The Impact of Early Environmental Conditions on the Evolution of Solar Type Stars As Seen In Cosmological Zoom-In Simulations of Milky Way Type Galaxies

by

Phoebe Stainton

A thesis submitted in partial fulfillment for the
degree of Doctor of Philosophy

January 2025

Declaration

The work presented in this thesis was carried out at the Astrophysics Research Institute, Liverpool John Moores University. Unless otherwise stated, it is the original work of the author.

While registered as a candidate for the degree of Doctor of Philosophy, for which submission is now made, the author has not been registered as a candidate for any other award. This thesis has not been submitted in whole, or in part, for any other degree.

Phoebe Stainton
Astrophysics Research Institute
Liverpool John Moores University
146 Brownlow Hill
Liverpool
L3 5RF
UK

January 2025

Abstract

The Sun is the only star known to host life, but, as of yet, there has been no clear reason as to why that should be the case. There are many other stars like the Sun, and exoplanet research has shown that there are thousands of other stars that host planets. The environmental condition of these planets, however, may be far less conducive to life than those found on Earth. To establish if this is the case, it is important to understand the early environmental conditions that the planet, and therefore host star, formed and evolved in.

Therefore, this research aims to quantify the impact of the cosmic environment on the formation and evolution on stars and their associated planets over cosmic time. Due to computational limitations, it is not currently possible to resolve individual stars or planet systems in cosmological simulations. Therefore, the stellar clusters found in these simulations must serve as proxies. This study utilizes Empirically Motivated Physics (EMP) simulations to investigate the relationship between particle metallicity and variables that may have impacted the formation and evolution of stars and planets.

Repeated trends are observed between particle metallicity and measured variables, suggesting that metallicity, often associated with regions of higher stellar density, plays a significant role in shaping the evolutionary history of stellar particles and, by extension, stars and their planets. We discuss each of these in detail and aim to determine how significant each of these may have been in determining the habitability of planets.

Acknowledgements

As is customary, I would like to thank my supervisors, Steve Longmore and Toby Moore, for their guidance and support over the last 3.5 years.

I would also like to thank the staff at the ARI for making my time there so enjoyable. Special thanks must go to my mentors, Ricardo Schiavon and Sebastian Kamann, for their encouragement, feedback, and guidance, especially in my final year.

To Tuts, Jemima, Andrew, Kyle, Adrian, Shobit, Sarah, Bethan, Flo, Stephen, Elena, Ryan, Andrea, and Sara, thank you for making 3.18 such a joy. Your coding support, paper recommendations, kitchen trips, agricultural skills, and contributions to the snack table have kept me going for the entirety of my PhD.

George, I think it's fair to say we met in some pretty extreme circumstances, but I think that is rather fitting. You are the most steadfast friend, and I am so grateful that you found me and stuck with me.

To Lisa, I have never met someone who is able to be so happy at such an unreasonably early hour. You never fail to make my day brighter, and I am truly grateful for your friendship.

I must of course thank my parents, Alan and Sally, and my brother, Harry. You have stuck by me and supported me in the most ridiculous of endeavours, and have always provided a safe haven for me to run to. You have been patient and kind, and have always guided me in the right direction. Most importantly, you have shown me that love really is all you need. Thank you.

To Spencer, you were by far and away the most opinionated member of the family, even though you couldn't talk. You brought me such joy and comfort, and I miss you immensely.

Finally - but most importantly - I want to thank my husband, Harris. I knew as soon as we met that it was the start of a grand adventure.

"I'm less "hot and ready" and more lukewarm and barely with it."

Dr. A.C.Mason

"As the days go by, we face the increasing inevitability that we are alone in a godless, uninhabited, hostile, and meaningless universe. Still, you've got to laugh, haven't you?"

Holly

Contents

List of Figures

List of Tables

Abbreviations

AD	**A**nderson-**D**arling
EMP	**E**mpirically **M**otivated **P**hysics
KS	**K**olmogorov-**S**mirnov

To those I have lost, I am better because you existed . . .

Chapter 1

Introduction

This thesis focuses on using cosmological simulations to determine if there are any links between the present day properties of stars, their evolutionary histories, and their potential for hosting habitable planets. This is a very broad piece of work which covers many areas of astrophysics that occur on a wide range of scales.

As this work aims to determine if habitability is impacted by the formation and evolutionary environments of stars, it is important to discuss habitability and the search for life. To this end, section 1.3 provides an overview of the history of the search for extraterrestrial intelligence and cosmic habitability. Section 1.4 discusses the formation and classification of galaxies and how the different formation and evolution histories of galaxies impacts the stellar populations within them. This is followed by section 1.5, which contains a discussion on the formation of low and high mass stars and how this provides evidence for the type of environment the Solar System was formed in. Finally, section 1.6 discusses the current theories surrounding planet formation and how this process is impacted by the local environment.

1.1 Definitions

Throughout this work, a range of language is employed that has the potential to cause confusion. Therefore, the following list provides definitions to frequently used words and their meanings in this context.

Cluster Environment: These are found within particles. Due to the resolution of the simulations, individual clusters cannot be identified, so there is no indication of their characteristics.

Gas Particle: Particles begin as gas particles. These may be converted into stellar particles under certain conditions.

Groups: The nine age-metallicity groups as defined in Section 3.2.

Halo: A structure containing a Milky-Way type galaxy and other particles that have become gravitationally bound to the galaxy.

Particle: a computational element used to represent a region of matter.

Run: The result of a completed simulation. In other papers (e.g. Reina-Campos et al., 2022a), these are referred to as haloes, but they have been renamed in this work to avoid confusion.

Stellar Particle: A representative of a stellar population. A particle may have mass in cluster environments and the field, but limits on the resolution of the simulation do not allow for these to be resolved into individual stellar clusters or stars.

1.2 Motivation

The evolutionary histories of stars help to shape their modern day observable properties. Certain properties, such as the number of supernovae experienced, will play a role in determining if a star is able to form planets, and if those planets are habitable. Currently, there is only one example of a star supporting life: the Sun. It is therefore important to question whether the early evolutionary conditions of sun-like stars differ from other types of stars, and whether these differences are more favorable to life.

In order to establish what modern day observational properties reveal about the evolutionary histories of stars, it is crucial that we are able to track the formation and evolution of stars. However, the time scales required for this work far exceed human life times, making this difficult to achieve through observational means. Therefore, simulations that model star and galaxy formation from the Big Bang to the present day are required.

The aim of this research is to quantify any differences or similarities in the early evolutionary histories of solar-type stars compared to stars with different ages and metallicities. By quantifying these, it may be possible to link modern day observable properties, such as age and metallicity, to the evolutionary history and the potential habitability of any planets that have been able to form.

Current work in this area is generally classified as the search for extraterrestrial intelligence (SETI). This is a broad term that encompasses a range of methods used to search

for life, including the direct detection of exoplanets, and the use of astrochemistry to study their atmospheres.

1.3 The History of SETI

Research into the habitability of planets is a small part SETI. Whilst curiosity surrounding life outside of that on Earth has been around for centuries, scientific searches for life are relatively new (Kellermann, 2023). In this section, a brief overview of the history of SETI is provided.

1.3.1 Early Searches

Early searches for extraterrestrial intelligence were confined to the Solar System. During his research in 1899, Nikola Tesla detected a signal he believed to be from Mars. In the early 1900s, Guglielmo Marconi (the inventor of the 'wireless telegraph'(Jabbari, 1997)), along with Lord Kelvin and David Peck Todd, supported Tesla's theory that radio could be used to contact any potential inhabitants of Mars.

Between 21-23 of August 1924, the United States implemented a National Radio Silence Day, with radios falling silent for 5 minutes every hour. This coincided with Mars entering its closest opposition to Earth in a century. Under the leadership of scientist David Peck Todd and Admiral Edward Eberle, a radio receiver at the US Naval Observatory was elevated 3km into the air and tuned to a wavelength of 8-9 km. If any extraterrestrial messages were received, US Army cryptographer William Friedman was in charge of translating it. Despite the best efforts of those involved, this experiment proved fruitless.

1.3.2 The Ohio State SETI Programme and the Wow! Signal

In 1955, John Kraus suggested that radio signals could be found using a flat-plane radio telescope with a parabolic reflector. By 1957, Ohio State University had approved the construction of what became known as 'Big Ear'. Following this, Ohio State began the first continuous SETI program.

On August 15th 1977, project volunteer Jerry Ehman noticed a strong signal picked up by Big Ear. It originated in the same area as the constellation Sagittarius, and was considered to have many of the expected properties of a signal from extraterrestrials. On the intensity printout, Ehman circled the signal and wrote 'Wow!' in the margin, giving

the signal its name. The signal is arguably one of the best candidates for extraterrestrial life, but it has not been detected again.

1.3.3 Frank Drake and the Drake Equation

In 1960, Cornell astronomer Frank Drake began work on Project Ozma, the first modern SETI experiment. Project Ozma utilised a 26m radio telescope at Green Bank, West Virginia to monitor Tau Ceti and Epsilion Eridani at the 1.42 GHz frequency. A 400 kHz band around this frequency was scanned with a single channel receiver, but nothing of interest was found.

Undeterred by the lack of results from Project Ozma, Drake decided to promote conversations around the potential for intelligent life within the Milky Way. To do this, he developed the Drake Equation:

$$N = R_* \cdot f_p \cdot n_e \cdot f_l \cdot f_i \cdot f_c \cdot L, \tag{1.1}$$

where R_* is the average rate of star formation in the Milky Way, f_p is the fraction of stars with planets, n_e is the average number of life supporting planets, f_l is the fraction of these planets that actually develop life, f_i is the fraction of planets with life that develop civilisations, f_c is the fraction of civilisations that develop technology that gives off detectable signals, and L is the length of time these signals are emitted for (Drake, 1961a).

As the final four parameters are unknown and have no reliable estimated values, solving the Drake equation is incredibly difficult. However, the equation serves to remind scientists of the range of factors that must be considered when searching for life elsewhere in the Universe via radio communication. In particular, the Drake equation has highlighted the lack of understanding around abiogenesis (the original formation of life from non-living matter), and has also helped to develop the field of astrobiology.

1.3.3.1 Estimated Values of the Variables in the Drake Equation

Despite there being very little opportunity to accurately constrain some of the variables in the Drake equation, efforts have still been made to provide estimated values.

Drake suggested a value for R_* of 1 yr^{-1} (Drake, 1961b), meaning 1 star would be formed per year in the galaxy, which was considered to be fairly conservative. The value of f_p was estimated to be between 0.2 and 0.5 (Drake, 1961b), meaning between a fifth

and a half of all stars will have formed at least one planet. The number of planets per star capable of developing life, n_e, was estimated to be between 1 and 5 (Drake, 1961b). Both f_l and f_i were estimated to be 1 (Drake, 1961b), meaning all planets capable of developing life would do so, and all these planets would develop intelligent life. The fraction of this life that will be able to communicate, f_c was estimated to be between 0.1 and 0.2, and the length of time that communications were detectable for, L, was given a value of 1000 - 100,000,000 years (Drake, 1961b). If the minimum estimates are used, this gives N = 20, but if the maximum values are used, N = 50,000,000.

Since Drake first published his equation, both technology and scientific understanding of the Universe has evolved. This has provided updated estimates for the variables in Drake's equation.

The value for R_* is now thought to be between 1.5-3 stars per year. This is based on research that suggests the star formation rate of the Milky Way is 0.68-1.45 $M_\odot$ per year (Robitaille & Whitney, 2010). If this is divided by the initial mass function for stars, the associated R_* value is 1.5-3 stars per year (Robitaille & Whitney, 2010).

Microlensing surveys within the Milky Way have highlighted that stars are almost always orbited by at least one planet (Cassan et al., 2012). This suggested that the value of f_p should be revised up from 0.2-0.5 to approximately 1.

The value for n_e, which represents the average number of planets that can support life, is much more difficult to constrain, although values between 0.1-1 are generally accepted. Kepler data suggests that, within the Milky Way, there could be up to 40 billion Earth-sized planets orbiting in the habitable zones of Sun-like and red dwarf stars (Petigura et al., 2013). As the Milky Way is home to approximately 100 billion stars, if all 40 billion stars hosted life, the value for n_e would be 0.4.

However, orbiting within the habitable zone does not necessarily mean that a planet would be capable of hosting life. Lineweaver et al. (2004), argued that, in order to host life, planets would need a suitable amount of heavy elements, a prolonged period of stability, and would also need to be far enough away from supernovae to avoid destruction. This work concluded that approximately 10% of star systems within the Milky Way could host life (Lineweaver et al., 2004), giving an n_e value of around 0.1.

The fraction of planets that actually go on to develop life, f_l is also subject to debate. Westby & Conselice (2020) suggests that the formation of life via abiogenesis is a natural part of the continued development of a habitable planet, and most planets will undergo this process. If this is the case, f_l should be close to 1.

However, this is countered by the argument that abiogenesis has only occurred once on Earth (Davies, 2007). If abiogenesis was common, it would likely have occurred more than once, and the lack of evidence for this suggests that this is a relatively uncommon process, giving an f_l value close to 0.

Another controversial value in the Drake equation is f_i, which denotes the fraction of life that develops into intelligent life. Some, such as biologist Ernst Mayr, argue that despite hosting billions of species, Earth has only developed one intelligent life form, humans (Mayr, 1995). This suggests that only a small fraction of life becomes intelligent (Mayr, 1995), implying that f_i should be close to 0.

The counter argument to this is that life generally becomes increasingly complex over time, suggesting that all life will eventually become intelligent (McKinney, 1989). This would suggest that f_i should be close to 1.

The next term, f_c represents the fraction of intelligent life that releases signals into space. This is also a speculative term, as there are many factors that need to be considered. For example, other civilisations may be releasing signals that current Earth technology cannot detect. Therefore, this value is practically unconstrained, and can vary between 0 and 1.

Finally, the amount of time a civilisation can communicate for, L, must be considered. Ćirković (2004) argues that L could potentially be of the order of millions of years. This value is dependent on the civilisation surviving large scale events, such as ecological collapse and hostile activity from other civilisations. Whilst an exact L value is not given, a range of 10^6 - 10^7 years is suggested under these conditions (Ćirković, 2004).

The argument for an large values of L has been countered by astronomer Carl Sagan and paleobiologist Olev Vinn. Both have posited that intelligent civilisations will eventually design technology that will lead to self-destruction (Vinn, 2024). This would prevent L from being very large, but does not limit the range of values beyond that.

1.3.4 Cosmic Habitability

The potential habitability of other planets has become a widely investigated topic in astrophysics. Within astrobiology, the characterisation of habitable planets within planetary systems outside our own has led to the discussion of the circumstellar habitable zone (CHZ) (Lineweaver & Chopra, 2012). This has been expanded on to account for the impact of the wider environment, leading to work on the galactic habitable zone (GHZ). This takes in to account the environments in which planets have formed, where

host stars have evolved, and cosmic events such as supernovae, which are considered in this work.

Supernovae are known to be crucial for synthesising heavy elements, such as aluminium and iron. These elements can also be found as a range of isotopes, which undergo radioactive decay due to their unstable nuclei. If this process occurs in the centre of forming planets, the high energy particles that are releasedby the decay of elements such as potassium, uranium, and thorium help to maintain the internal temperature (Dye, 2012). This central process is usually dominated by isotopes with short half lives, and isotopes with longer half lives are usually found nearer the surface of the planet, where their decays help to maintain the surface temperature (Kaib, 2018).

However, the energy released in a supernova has the potential to damage both forming and established planets (Ellis & Schramm, 1995). The EUV and XUV produced is capable of destroying protoplanetary disks, preventing planets from forming. If a star already hosts established planets, the energy can instead strip the atmosphere from the planets, negatively impacting the habitability.

Stars in cluster environments are subjected to a higher stellar density (Kopparapu, 2005), which increases the likelihood that a nearby star will go supernova. As previously mentioned, supernovae release large amounts of energy, which are capable of disrupting protoplanetary disks or the atmospheres of established planets (Kaib, 2018). Combined with the increased risk of gravitational interactions, this makes cluster environments more hostile to life than field environments.

When stars are in cluster environments, they also have lower relative velocities compared to field stars (Kopparapu, 2005), which, along with the increased stellar density, will increase the likelihood of gravitational encounters between stars (Kaib, 2018). These encounters could lead to planetary orbits being perturbed, or even to planets being ejected (Laughlin & Adams, 1998; Spurzem et al., 2009).

Recent observational work (e.g Madhusudhan et al., 2023; Ahrer et al., 2022) highlights the detection of molecules such as carbon dioxide and dimethyl sulfide in the atmospheres of exoplanets. As both of these molecules are produced in processes associated with life (Keller, 1989), they are considered to be good indicators of liquid water (Keller, 1989), which is known to be crucial to life.

Observational work is ongoing (e.g Tsumura, 2020; Enya et al., 2023, and references therein), and includes work on detecting new exoplanets, the study of atmospheres, and the constraining of measurements from previously identified planets. However, it is an expensive process which benefits from the use of modelling to identify stars that are likely to host habitable planets. As this work aims to identify the stars statistically most

likely to host habitable planets, it is possible that this work could be used to highlight good targets for future observations.

1.3.5 This Work in the Context of SETI

The overarching aim of this work is to establish how evolutionary environments, such as the amount of time spent in a cluster environment and the number of supernovae experienced, impact stellar evolution and the potential for planetary development. With additional work looking at the potential habitability of these planets, this work would help to provide a limit for the value of n_e in Equation 1.1. Additionally, by determining how present day stellar properties are impacted by their evolutionary environments, our work can be used to determine if some stars have been in such hostile environments that they are unlikely to have formed any planets that have survived. This could then be used to rule these stars out stars from observation runs. For any target stars that remain, our work can be used to prioritise them based on the likelihood that they host planets.

1.4 Galaxy Classification and Formation

1.4.1 Galaxy Classification

Galaxies are dynamically bound systems that host a stellar mass between $10^5 M_\odot$ (Kirby et al., 2013) to $10^{11} M_\odot$ (Baldry et al., 2012).

Galaxies are often classified based on their morphology. The most common system used is the Hubble sequence (or tuning fork). The galaxies are classified as elliptical, spiral, or irregular galaxies, with some groups being split into sub groups.

Elliptical galaxies are usually denoted as EX, with X being an integer ranging from 0 to 7. The number denotes how spherical the galaxy is, with 0 being the most spherical and 7 being the least. Star formation in these galaxies is thought to have occurred billions of years ago during a burst of star formation activity. This star formation activity is no longer occurring (De Lucia et al., 2006), although gas accretion and mergers may induce a small amount of activity. This means that the majority of stars in elliptical galaxies are old, low mass, low metallicity stars formed in the early universe (De Lucia et al., 2006). These stellar populations give elliptical galaxies their 'redder' photometric colours (Sellwood & Masters, 2022).

Spiral galaxies, like the Milky Way, are characterised by their spiral arms and flat disks (Sellwood & Masters, 2022). These galaxies are then split into sub groups based on the characteristics of their spiral arms.

The first sub group are described as normal spirals, and are denoted as Sa, Sb, and Sc. The spiral arms in these galaxies are well defined (Sellwood & Masters, 2022). Sa galaxies are identified by their tightly wound arms and a large stellar bulge, with the tightness of the arms and size of the bulge decreases through the remaining two groups (Sellwood & Masters, 2022). The second group is barred spirals, which are denoted as SBa, SBb, SBc. These galaxies display the same traits with regards to their spiral arms, but the all have a bar structure running through their centre which is connected to the arms (Sellwood & Masters, 2022).

All spiral galaxies have ongoing star formation, which is often enhanced in their arms and disks (Sellwood & Masters, 2022). Barred spiral galaxies also have a region of enhanced star formation inside their bars. Due to this ongoing star formation, the stars found in spiral and barred spirals have a range of ages and metallicities (Bell & de Jong, 2000). Photometrically, spiral and barred spiral galaxies are bluer than their elliptical counterparts (Bell & de Jong, 2000). However, as certain regions of these galaxies can have different star formation histories, the photometric colour can vary across a given galaxy (Bell & de Jong, 2000).

The final group are irregular galaxies. As the name suggests, these galaxies do not fit into either of the groups described above due to their lack of structure (Hunter, 1997). Irregulars are divided into 2 sub categories: galaxies with some poorly defined structure (Irr I), and galaxies with no obvious structure (Irr II).

Irregular galaxies tend to have either ongoing star formation, or display signs of recent star formation (Hunter, 1997). However, there is minimal uniformity in their star formation histories beyond this, with some irregulars undergoing continuous star formation, and others having quiescent periods punctuated by periods of intense star formation that can be triggered by mergers or other interactions (Hunter, 1997).

Due to the nature of their star formation histories, irregular galaxies play host to a range of stars. There are often some older, redder stars in areas of the galaxy that have not undergone star formation recently. However, irregular galaxies tend to have a large number of young, blue stars due to their recent or ongoing star formation (Hunter, 1997).

1.4.2 Galaxy Formation

Initial galaxy formation models consisted predominantly of 'top- down' theories, with the most famous being the monolithic theory of Eggen et al. (1962), which is based on observations that show a correlation between the metallicity, orbital eccentricity, and orbital angular momentum of stars in the solar neighbourhood (Eggen et al., 1962).

To explain the link between these properties, it was concluded that the Milky Way was formed when a proto-galactic nebula collapsed (Eggen et al., 1962). They suggest that the oldest halo stars were formed near the beginning of the collapse, leaving them with more radial trajectories and highly elliptic orbits that lie above and below the galactic plane (Eggen et al., 1962). As these stars formed very early, heavy elements usually formed in nucleosynthesis (Woosley & Weaver, 1995) were lacking, making these early stars metal poor (Pop II). As massive stars began to form, synthesise heavy elements, and go supernova, material containing heavy elements was ejected into the ISM (interstellar medium) (Woosley & Weaver, 1995), enriching it.

As the collapse process continued, the radius of the proto-galactic cloud decreased (Eggen et al., 1962), increasing the frequency of collisions between dust and gas. These collisions converted kinetic energy into thermal energy and slowed the rate of collapse. The decrease in radius and conservation of angular momentum caused the collapsing cloud to rotate more quickly, eventually forming a chemically enriched disk that was able to form Pop I stars.

However, Searle & Zinn (1978) observed that red giants in globular clusters found in the galactic halo had a range of ages and metallicities. If galaxy formation had occurred via a monolithic mechanism, the stars in the globular clusters should not display this range, and should instead have similar ages and metallicities.

The possibility of dark matter was also being probed at this time. The consequences of dark matter had been observed for decades, but the importance of these results had not been fully understood or accounted for. Kapteyn studied the distribution of stars in the Milky Way, and provided a description of the distribution of masses and velocities of these stars. He described the Milky Way as being flat, and observed that it rotated around an axis (Kapteyn, 1922). He then established a relationship between the velocity dispersion and motion of stars in the Milky Way (Kapteyn, 1922).

In 1933, Zwicky studied the redshifts of galaxy clusters and noted an unexpectedly large scatter in the recorded apparent velocities of eight galaxies in the Coma Cluster (Zwicky, 1933). Zwicky began by applying the virial theorem to estimate the mass of the Coma Cluster, and then calculated the average kinetic energy and velocity dispersion. His

calculated velocity dispersion of 80 km/s was significantly smaller than the observed value of 1000 km/s (Zwicky, 1933). This suggested that there was more matter that was not being accounted for.

Zwicky's work was expanded upon in 1939, when PhD student Horace Babcock measured the rotation curve of M31 from its centre to 100 arcminutes (roughly 20 kpc) (Babcock, 1939). The resulting rotation curve showed an increase in orbital velocities at large radii, suggest there was a large amount of undetected mass in the outer regions of M31 (Babcock, 1939).

By the 1960s and 70s, technology had advanced sufficiently that more detailed measurements could be made of galaxies and their rotation curves. Freeman used these measurements to compare the predicted peaks of rotation curves to the observed 21cm curves (Freeman, 1970). The predicted rotation curves were based on the assumption of an exponential disk and utilised scale heights fit to photometric data (Freeman, 1970). His results showed that the rotation curves for M33 and NGC300 peaked at larger radii than predicted, suggesting that there was additional mass in these galaxies (Freeman, 1970).

This result was supported by the work of Rubin and Ford, who had published spectroscopic observations of M31 (Rubin & Ford, 1970). They later collaborated with Thonnard, and published optical rotation curves for ten high-luminosity spiral galaxies (Rubin et al., 1978). Both pieces of work showed that the rotation curves peaked before becoming roughly stable, supporting the theory that there was additional mass in these galaxies.

Combined, this work suggested that dark matter made up the vast majority of mass in galaxies. As this was not accounted for in the monolithic theory, new theories, such as that of White & Rees (1978), were developed to take this into account. Eventually, these 'bottom-up' (or hierarchical) theories replaced the monolithic theory.

The most widely accepted hierarchical model is based on the cold dark matter (ΛCDM) model, where Λ is the cosmological constant. This suggests that large structures are the result of the accretion of smaller structures. This process begins when localised regions of dark matter collapse and increases in density. This is mimicked throughout space, leading to the formation of dark matter haloes. These haloes have a strong gravitational potential, which pulls in any surrounding baryonic matter, forming larger structures, including galaxies. Spiral galaxies are thought to have formed after very few interactions, whereas their elliptical counterparts are likely to have required more mergers to form (Bournaud et al., 2007).

Recent work suggests that galaxy evolution is an ongoing, hierarchical process, with growth occurring via both in situ star formation and satellite accretion (Kruijssen et al., 2019b). This theory is supported by recent observations, which have shown matter from tidally disrupted dwarf galaxies is present in the halo of the corresponding central galaxy (Belokurov & Kravtsov, 2022; Deason & Belokurov, 2024).

1.5 Star Formation and Star Formation Efficiency

For galaxies to exist, it is crucial that stars are formed. Star formation is a complex process, with many factors influencing it. As star formation is not the main focus of this work, this section aims to present an overview of star formation.

1.5.1 Molecular Clouds, Cores, and Star Formation Efficiency

Within galaxies, stars are born in molecular clouds. These clouds have temperatures of around 10-20K, and densities of approximately $n_H \approx 10^{2-5}$ cm^{-3}. These clouds are not homogeneous, and instead have a range of substructure, including 'clumps' (Elmegreen, 2008). Clumps are on the scale of $0.1 - 1$ pc, and within these clumps are even more compact overdensities called cores, which exist on scales $\ll 0.1$pc (McKee & Ostriker, 2007).

Star formation occurs when clouds and, by extension, the dense core embedded within them, collapse under gravity. The cloud will collapse over a free-fall time, which is given by:

$$t_{ff} = \left(\frac{\pi^2 R_c^3}{8GM_c}\right)^{\frac{1}{2}} = \left(\frac{3\pi}{32G\rho_c}\right)^{\frac{1}{2}}, \tag{1.2}$$

where R_c is the cloud radius, M_c is the mass of the cloud, G is the gravitational constant, and ρ_c is the average density of the cloud. This gives free-fall times on the order of a few Myrs (Chevance et al., 2020). However, if all the gas in the Milky Way collapsed to form stars in that time frame, the observed star formation rate would be around 2 orders of magnitude higher than is observed. It therefore follows that only some of the cloud mass, M_{gas} is converted into stellar mass, $M_{*,tot}$. The fraction that is converted into stellar mass is the star formation efficiency, ϵ, and is given by:

$$\epsilon = \frac{M_{*,tot}}{M_{*,tot} + M_{gas}} \tag{1.3}$$

As the observed order of magnitude for star formation rate is so much lower than what would be expected if all the gas was converted to stellar mass within one free-fall time, it follows that star formation is an incredibly inefficient process (McKee & Ostriker, 2007). To account for this in a wider astrophysical context, several mechanisms that inhibit cloud collapse have been proposed. The simplest of these is the inclusion of the virial parameter, α_{vir}. In the simplest example, we consider a spherical cloud with a uniform density that is supported solely by kinetic energy. The virial parameter can then be expressed as (Jeans, 1902):

$$\alpha_{vir} = \frac{5\sigma_{\nu}^2 R_c}{GM_c}, \tag{1.4}$$

where σ_ν is the velocity dispersion, G is the gravitational constant, and M_c and R_C are the mass and radius of the cloud respectively. If $\alpha_{vir} > 2$, the cloud is said to be 'super-virial' and will expand. If $\alpha_{vir} < 2$, the cloud is 'sub-virial' and will collapse. If $\alpha_{vir} = 2$, the cloud is stable. There is thought to be a link between α_{vir} and the star formation efficiency per free-fall time, but this link is not well defined and remains controversial (Schinnerer & Leroy, 2024).

1.5.2 Formation of Low Mass Stars

Along with the SFE, the formation of stars is dependent on the stellar initial mass function (IMF). An IMF is a probability distribution of the birth masses of stars. Due to the physical processes that occur in star formation, cloud collapse, and conservation laws, IMFs tend to predict a larger number of lower mass ($< 8M_\odot$) stars. Low mass stars are classified as being below $8M_\odot$ as these stars do not meet the mass threshold to become a type II supernova (Zinnecker & Yorke, 2007), meaning they have a different evolutionary history to their high mass counterparts.

One important way in which these stars vary is in their Kelvin-Helmholtz timescale. This is the defined as the approximate time it would take for a star at its current luminosity to radiate away all of its gravitational potential energy, and it is given by:

$$t_{KH} = \frac{1}{2}\frac{E_{GPE}}{L_*} \approx \frac{GM_*^2}{R_* L_*}, \tag{1.5}$$

where E_{GPE} is the gravitational potential energy, L_* is the luminosity of the star, G is the gravitational constant, M_* is the mass of the star, and R_* is the radius of the star. For the Sun, this is approximately 31 Myrs, and for Sun-like stars, this is of the order of 10 Myrs. These timescales are significantly longer than that of pre-main sequence

accretion, which occurs on timescales between hours and centuries (Hartmann et al., 2016). This means that low mass stars accrete all of their mass before they enter the zero age main sequence. However, as L is proportional to $M^{3.5}$, the t_{KH} is very short got high mass stars, meaning they do not have a similar 'pre-main sequence' phase.

Another crucial difference is the level of understanding of the formation of both types of star. The nature of the IMF means that more low mass stars have formed in our local environment, allowing for greater study of their evolutionary paths. It is therefore well known that low mass stars begin their formation when protostellar cores collapse (McKee & Ostriker, 2007).

In the initial stages of the collapse, the dust is optically thin, meaning that the radiation produced by the collapsing core is able to escape, leaving the core with a roughly stable temperature, but a decreasing radius and increasing density. This constant temperature means the core can be treated as having a uniform temperature (i.e it is isothermal). As the collapse continues, the density increases and the dust eventually becomes optically thick, leading to an increase in the temperature (Larson, 1969). As the temperature increases in the core, the subsequent pressure increase prevents collapse and the core is said to be hydrostatic. This usually occurs when the central density has reached $\approx 10^{-10}$ g cm^{-3}. However, the temperature in the core continues to increase, and at $\approx$ 2000K molecular hydrogen begins to split into separate atoms. At this point, the heat being generated by the compression of the core is used to disassociate molecular hydrogen instead of providing the pressure needed to prevent collapse. This allows a second collapse to occur and the core contracts isothermally. When the majority of the hydrogen in the core has become ionised, the contraction stops and the core returns to a hydrostatic state. This leaves a proto-star, which continues to grow via accretion before entering the zero age main sequence.

1.5.3 Formation of High Mass Stars

High mass stars are much rarer than their low mass counterparts due to the shape of the IMF. However, high mass stars are crucial for the synthesis of heavy elements that occurs during their evolution. At the end of their evolutionary paths, these stars will go supernova, ejecting these heavy elements into the ISM (Ott, 2016).

Despite the important role they play in the evolution of star and galaxies, the formation of high mass stars is poorly understood. This is in part due to their rarity, but is also impacted by their relatively short lifetimes. There are currently 2 main theories under consideration for the formation of these stars: core accretion and competitive accretion (Tan et al., 2014).

The core accretion model merely extends the theory used for low mass stars, assuming that the formation of high mass stars occurs when cores of gas collapse under gravity.

The theory is extended in part by including optically thick accretion disks, which enable material to accrete onto the protostar (Krumholz, 2015). These disks are dense, thick, and unstable, meaning they can form spiral structures that channel gas more efficiently than a normal disk (Kratter & Matzner, 2006). This efficiency causes an increase in the accretion rate, which is required to form higher mass stars.

Another critical difference that must be accounted for is the increase in turbulence. This increased turbulence prevents the molecular cloud from collapsing too early, allowing massive cores to form in local overdensities (McKee & Tan, 2003). This allows fewer, more massive stars to form instead of many smaller stars.

In the competitive accretion model, high mass stars are not formed in a gravitationally bound core. Instead, gas is drawn from the wider clump that contains the cores, before collapsing to form a high mass star (Tan et al., 2014). This theory therefore suggests that high mass stars are regularly surrounded by low mass stars. In cases where there are a significant number of low mass stars surrounding a high mass star, stellar collisions occur, increasing the size of the high mass star (Tan et al., 2014).

1.5.4 The Solar System and Star Formation

The formation of stars of different masses plays an important role in our understanding of the formation of our solar system. As star formation theories have evolved, so have the theories governing the formation of the solar system.

Early theories of the formation of the solar system suggested that the process occurred in isolation, with no other stars nearby. This theory is generally supported by observations of nearby star and planet forming regions, but more recent evidence suggests that this theory does not hold true in all cases.

The isolated nature of this process is brought into doubt partly due to the existence of radioactive isotopes in the solar system.

Lee et al. (1976) highlighted the exceptionally high amount of ^{26}Mg in the Allende meteorite. Further work by Davis et al. (2014) found that the early solar system was likely home to large amounts of ^{26}Al, the decaying of which would explain the large amounts of ^{26}Mg found by Lee et al. (1976).

A potential source of the ^{26}Al is the dust in the solar protoplanetary disk. If the disk was exposed to large amounts of radiation, ^{26}Al could have been synthesised from the dust.

However, Gounelle et al. (2013) found that this was unlikely, meaning that the isotope must have been formed via nucleosynthesis. As the stellar synthesis of ^{26}Al requires high temperatures, this process must have occurred in a star larger than $8M_{\odot}$ (Davis et al., 2014).

This suggests that the ^{26}Al found in the solar system is the product of at least one large, nearby star that has gone supernova, enriching the ISM (Ott, 2016). However, a more likely scenario is that the solar system formed in a stellar cluster, with many stars contributing to the ^{26}Al abundance.

1.5.5 Star Clusters

Lada & Lada (2003) suggests that stars regularly form in cluster environments. These clusters vary in mass and age, but can contribute a significant number of stars to their host galaxy (Goddard et al. (2010); Kruijssen (2012a)). Originally, clusters were split into one of two groups: globular clusters, which are old ($\gtrsim$ 10 Gyr), have high masses ($\gtrsim 10^5\ M_{\odot}$), and high densities ($\gtrsim 10^3\ M_{\odot}$ pc^{-3}), and open clusters, which are younger ($\lesssim$ 1 Gyr), less massive ($\lesssim 10^3\ M_{\odot}$), and less dense ($\lesssim 10^3\ M_{\odot}$ pc^{-3}). Globular clusters are also required to be gravitationally bound, whereas open clusters do not need to be. However, more recent work (e.g. Portegies Zwart et al., 2010) has defined young massive clusters (YMCs) as having masses on the order of $10^4\ M_{\odot}$ and ages $\lesssim$ 100 Myrs. YMCs are gravitationally bound and still forming (Holtzman et al., 1992).

YMCs have been found to host lots of high mass stars alongside large numbers of low mass stars. These dense environments would enable the formation of high mass stars through competitive accretion, and would also enable the chemical enrichment of planetary systems via nearby supernova, such as the solar system.

1.6 Planet Formation

After stars have formed, they may go on to form planets, although this process is highly dependent on the environment stars inhabit. The nature of planet formation is complicated, so this section only provided a brief overview. For more details, I direct the reader to Raymond & Morbidelli (2022) and Drazkowska et al. (2023) and references therein.

1.6.1 The Solar Nebular Disk Hypothesis

One of the first widely accepted models of planetary formation was the nebular hypothesis (Kant, 1755). This hypothesis was based on observations of the Solar System, and suggested that it was formed from gas and dust that had clumped together whilst orbiting the Sun (Kant, 1755). This model was, generally speaking, correct, and is now known to be applicable to other planetary systems (Süli, 2010). The modern version of this theory is known as the solar nebular disk model (SNDM) (Süli, 2010). The basis of the SNDM is that the early universe contained a gravitationally unstable gas that formed many dense, rotating clumps that collapsed into flat disks in order to conserve angular momentum (Clement et al., 2024). These disks surround stars and are capable of forming planets (Clement et al., 2024).

The validity of the SNDM is easily tested with observations of our Solar System. If the SNDM is correct, some predictions can be made about observable properties of the planetary systems formed:

1. The planets would need to have small inclinations relative to their host star.

2. The planets and their host star will have all formed from the same rotating matter, so the planets should orbit their host star in the same direction as the star itself rotates.

3. The orbits of the planets would need to be roughly circular.

In the Solar System, we see planets orbiting the Sun in the same direction as it rotates. Their orbits are coplanar and approximately circular, supporting the SNDM.

1.6.2 Early Stages of Planet Formation

A star similar to the Sun takes around 1 million years to form (Montmerle et al., 2006). The associated protoplanetary disk will form a planetary system over the following 10-100 million years (Montmerle et al., 2006). When the disk is in the T Tauri stage, small dust grains form in the disk. As these grains travel around the disk, they are prone to undergoing collisions that causes the grains to bind together. In smaller disks, there is not enough material to form anything larger than a planetesimal. In larger disks, it is possible to form larger bodies, but the composition of these bodies is dependent on the distance from the planetesmials and the host star.

1.6.3 Formation of Rocky Planets

The SNDM suggests that rocky planets will form within the frost line, which is found at around 3-4 AU from Sun-like stars (Montmerle et al., 2006). Within the frost line, the temperature is high enough that ice grains cannot form (Raymond et al., 2007). This means that planetesimals closer to the host star are only formed of rock (Raymond et al., 2007; Kley, 2019). Once the planetesimals have a diameter on the order of 1 km, runaway accretion can begin (Kokubo & Ida, 2002). This process lasts between 10,000 and 100,000 years, and will end when diameters of the larger bodies (oligarchs) is around 1000 km (Kokubo & Ida, 2002; Kley, 2019).

At this stage, the oligarchs dominate the accretion process, preventing any other bodies from growing (Thommes et al., 2003). These oligarchs continue to accrete planetesimals until there are none left in their local environments (Kokubo & Ida, 2002; Kley, 2019). This stage of rocky planet formation lasts around 200,000 years (Kokubo & Ida, 2002), and ends with roughly 100 bodies with sizes ranging from moon sized to Mars sized (Raymond et al., 2006).

Finally, the remaining bodies begin to collide with each other at low velocities, forming fewer, larger bodies with sizes similar to Earth. These new planets are not initially stable, but they will eventually find themselves in stable orbits around their host star (Raymond et al., 2006).

1.6.4 Formation of Giant Planets

The formation of giant planets is less well understood. It is known that this process occurs beyond the frost line (approximately 4AU for stars similar to the Sun), which means that the grains are composed predominantly of ice (Inaba et al., 2003). The most commonly accepted theory of giant planet formation is the core accretion model (or nucleated instability model) (D'Angelo et al., 2010).

This process is initially similar to the formation of rocky planets. The first stage is the accretion of dust into planetesimals and oligarchs to form bodies with masses of 10 M_E in approximately 1 million years (Thommes et al., 2003). At this stage, the bodies have a high enough gravitational pull to begin accreting gas (Pollack et al., 1996; Kley, 2019). This continues until the mass of the body reaches 20-30 M_E (Pollack et al., 1996). At these masses, the gas will gravitationally collapse on to the planet, prompting the start of runaway growth that sees the planet reach masses several hundred times that of Earth (Pollack et al., 1996; Kley, 2019). This occurs in approximately 10,000 years (Pollack et al., 1996).

Whilst this theory works well for planets such as Jupiter and Saturn, it does not account for the formation of Uranus or Neptune, as these planets are in regions where the matter density would be too low for this process to occur (Kley, 2019). One potential solution is that Uranus and Neptune formed in the same way as Jupiter and Saturn before they migrated further out (Thommes et al., 1999).

A second suggestion is that the cores of giant planets grew via pebble accretion rather than exponential gas accretion. In this theory, pebbles with diameters between 1cm-1m fell towards the cores of giant planets (Lambrechts & Johansen, 2012). Drag from the gas surrounding these cores slow down the pebbles, allowing the gravitational force of the core to accrete the pebbles (Lambrechts & Johansen, 2012). This process is significantly quicker than growth via planetesimal accretion (Lambrechts & Johansen, 2012), allowing the cores to begin the accretion of gas earlier. However, Uranus and Neptune would not grow to the size of Saturn or Jupiter due to starting gas accretion much later, when much of the gas had already been accreted (Papaloizou et al., 2006).

Based on these works, it is clear that the formation of both rocky and giant planets is highly sensitive to the environment the host star is found in. Therefore, the types of planets formed and their potential habitability will be partially determined by the evolutionary history of the host star and its environment.

1.6.5 The Impact of Photoevaporation and Metallicity on Planet Formation

The photoevaporation of protoplanetary disks due to UV and X- ray radiation poses a significant problem for planet formation. When radiation is incident on the disk, the outer layers of material are heated to approximately 10^3 - 10^4K. At a certain temperature dependent radius from the disk centre, this material becomes an unbound photoevaporative wind. The radius at which the occurs can be calculated via:

$$R_c \simeq \frac{0.2GM_*}{c_{hot}^2} \simeq 1.8\frac{M_*}{1M\odot}\left(\frac{T_{hot}}{10^4K}\right)^{-1} AU \tag{1.6}$$

where M_* is the stellar mass of the star in the disk, c_{hot} is the speed of sound in the heated disk layer, and T_{hot} is the temperature of the heated layer (Alexander, 2014; Font et al., 2004).

The heating of this material and its subsequent loss reduces the amount of material available to form planets. This may mean that fewer planets are formed in the disk, and those that do form may be smaller.

The loss of material may also dramatically alter the chemical composition of the disk. Swain et al. (2024) found that gas giants were more likely to form in regions of higher metallicity, whilst terrestrial planets were formed in lower metallicity environments. This suggests that if the winds reduce the metallicity of the disk, the planets formed are more likely to be rocky, with gas giants more likely to be formed if the metallicity increases (Swain et al., 2024).

The findings of Swain et al. (2024) are not surprising. Work by Fischer & Valenti (2005) found that whilst less than 3% of stars with a subsolar metallicity had detected planets, 25% of stars with an [Fe/H] value greater than 0.3 dex hosted gas giants. Metal rich stars were also found to be more likely to host multi-planet systems (Fischer & Valenti, 2005). Whilst it is possible that these stars increased their metallicity values through late-stage accretion, this process would leave detectable traces, such as an increase in metallicity as the mass of the convective envelope decreases (Fischer & Valenti, 2005). As there were no traces of late- stage accretion, the high metallicity seen in these stars is likely an inherited characteristic from the molecular cloud that formed them (Fischer & Valenti, 2005).

1.7 This Work

This work utilises simulations to track the formation and evolution of star particles over cosmic time, with a particular focus on stars like the Sun as this star is known to host life. With this in mind, it is crucial that this work considers what has already been done in the search for extraterrestrial intelligence, and what is known to be an indicator or requirement for life. It is also important to consider where different stellar populations are found, and if those environments are particularly hostile to stellar and planetary evolution. It is therefore necessary to consider the physical processes that govern the formation and evolution of stars and planets and how these processes may be impacted by the environment, and what this means for any potential life.

Chapter 2

Cosmological Simulations

2.1 The Use of Cosmological Simulations for This Work

In order to determine if the evolutionary environments of stars impact their ability to host life, it is necessary to track the formation and evolution of thousands of stars over cosmic time. If this was done observationally, it would require large amounts of observing time for a range of target stars, and many generations of astronomers.

Cosmological simulations provide a way to avoid these problems. The physics used in these simulations has advanced greatly in the last 15-20 years (Crain & van de Voort, 2023), with models now able to include cluster formation and evolution, and the physics of the cold ISM. Whilst this increased complexity does require significant amounts of computational time, the results obtained match well with observational results. This high level of accuracy makes simulations an ideal tool for modelling large stellar populations over cosmic time, hence their use in this work.

2.2 Background to Hydrodynamical Cosmological Simulations

Cosmological hydrodynamical simulations are designed to evolve both dark matter and baryons in a self-consistent, simultaneous manner. The evolution of dark matter and baryons is modelled at selected points during the simulation, usually known as snapshots. The time between snapshots is chosen based on the required level of precision and the available computing time, and it is not required to be uniform throughout. However, the time steps are often selected based on timescales relevant to the processes being modelled to ensure that important information is captured. At each snapshot, data for

each monitored variable of each particle is taken, allowing the evolution of individual particles to be followed over cosmic time.

2.3 Subgrid Models

As the capacity for resolution is limited, some processes, such as star formation and stellar feedback methods, are modelled in a semi- analytic way, meaning analytical models and numerical simulations have been used. These are referred to as subgrid models. The scale over which these models are utilised depends largely on the overall scale of the simulation.

The input parameters used by subgrid routines are tuneable, meaning they can be modified based on observational values. This allows simulation suites to be updated as more observational information becomes available, and also allows simulations to utilise a range of different values for constants that are less clearly defined. The validity of these input values is usually judged based on the ability of the simulation to accurately reproduce observed characteristics of the Universe, such as the galaxy stellar mass function.

2.4 EAGLE

Arguably the most well known suite of cosmological hydrodynamical simulations is EAGLE (Evolution and Assembly of GaLaxies and their Environments). This is a suite of galaxy formation simulations that use ΛCDM cosmogony (Crain et al., 2015; Schaye et al., 2015) and are evolved using GADGET3 (Springel, 2005), which incorporates subgrid routines that phenomenologically model physical processes occurring on scales smaller than the resolution. The initial models are calibrated to observations, and utilise volumes with sides of 25 to 100 comoving Mpc (cMpc). Comoving distances are used as this coordinate system expands with the universe over time. This means that the distance remains fixed whilst still accounting for the expansion of the Universe.

The resolution allows the Jeans scale in the warm ($T \sim 10^4$K) interstellar medium to be resolved (Schaye et al., 2015). The Jeans scale is the length scale below which gravitational collapse can occur, meaning that processes involving gravitational collapse (e.g. structure formation) can be modelled in EAGLE. The Jeans length is also used as the density seen in self gravitating systems (such as collapsing gas clouds) varies on this scale.

The temperature limit of 10^4K is imposed in EAGLE for several reasons. The first is that temperatures below this are usually of importance on small scales (e.g. star

formation). As EAGLE cannot resolve down to this scale, these temperatures are not modelled. Secondly, at these temperatures there are many non-linear processes at play (e.g magnetohydrodynamic turbulence). The non-linear nature of these processes mean that they would need to be modelled using smaller time steps to ensure adequate accuracy. This would increase the amount of time required for the simulation to run, which in turn increases the computing cost. For these reasons, 10^4K is implemented as the temperature floor in EAGLE.

Gas particles in EAGLE are able to convert into a stellar particle when they exceed the density threshold for star formation (Schaye, 2004). This threshold is metallicity dependent (Schaye, 2004), and has a value of 0.1 n_H cm^{-3}. Exceeding the threshold does not guarantee that a gas particle will be converted to a stellar particle. This threshold prevents high density gas from fragmenting into smaller clumps that cannot be resolved by the simulation, and also prevents high metallicity, low density particles from over cooling and entering a temperature regime that is not modelled. The probability of a particle undergoing a stochastic conversion depends on the gas pressure of the particle (Schaye & Dalla Vecchia, 2008), as particles with too high a gas pressure would be unable to form stars. Above the minimum threshold density, the star formation rate is expressed as a pressure law and calculated as:

$$\dot{m} = m_g \times A \times (1 M_{\odot} pc^2)^{-n} \times \left(\frac{\gamma}{G_N} f_g P \right)^{\frac{n-1}{2}} , \tag{2.1}$$

where n is the exponent of the Kennicutt-Schmidt relation (usually 1.4), A is the normalisation factor (usually 1.515 $\times$ 10^{-4} M$_{\odot}$ yr^{-1} kpc^{-2} for a Chabrier IMF), m$_g$ is the gas particle mass, γ is the adiabatic index, f$_g$ is the gas fraction of the disk, and P is the total pressure of the gas.

Using this star formation rate, the probability is then calculated as:

$$Prob = min \left(\frac{\dot{m} \Delta t}{m_g}, 1 \right) . \tag{2.2}$$

A random number is then generated, and if this number is greater than the probability, the gas particle undergoes a conversion.

To describe radiative cooling and photoionisation, the method outlined in Wiersma et al. (2009a) is utilised. Radiative cooling is necessary to ensure that thermal energy is dissipated, thus maintaining thermal equilibrium and allowing accreted gas in galaxies and molecular clouds to form stars. Photoionisation is the process by which photons ionise atoms or molecules, and is crucial for preventing excess star formation by ionising gas and preventing its collapse. The method used applies to hydrogen, helium, carbon,

nitrogen, oxygen, neon, magnesium, silicon, sulphur, calcium, and iron (Wiersma et al., 2009a), as these elements have the greatest impact on the cooling rates. The radiation field is taken to be spatially uniform and temporally-evolving (Crain et al., 2015), which is a valid assumption on large spatial scales with low spatial resolutions, as the effects of small scale variations will not be resolved. The gas is assumed to be optically thin and in equilibrium with respect to any ionisation (Crain et al., 2015), which is valid at low densities. If the gas has a density larger than the metallicity- dependent threshold described in Schaye (2004), it can be stochastically converted into a collisionless stellar particle. The probability of a conversion occurring is proportional to the star formation rate (SFR) of the particle (Crain et al., 2015). Each particle represents a simple stellar population that has an initial mass function (IMF) described by Chabrier (2003). Mass and metals that return from the stellar populations to the ISM are modelled by Wiersma et al. (2009b), which follows the same 11 species as the model from Wiersma et al. (2009a).

If a halo is above 10^{10} $M_{\odot}$/h (where h = $H_0/100 kms^{-1}$ Mpc^{-1} = 0.6777) and does not already contain a black hole, a seed black hole is placed at the centre of the halo (Springel, 2005). In these cases, the highest density gas particle is converted into a collisionless black hole particle rather than a stellar particle. The newly seeded black hole inherits the mass of the original gas particle and gains a subgrid black hole mass of 10^5 $M_{\odot}$/h.

The growth of these black holes is modelled via subgrid routines that govern black hole growth through gas accretion (which cannot exceed the Eddington limit) and black hole-black hole mergers (e.g. Schaye et al., 2015). There are also routines to model feedback from star formation (Dalla Vecchia & Schaye, 2012) and AGN (active galactic nuclei) feedback (e.g. Schaye et al., 2015), via stochastic gas heating (Dalla Vecchia & Schaye, 2012). The efficiency of the stellar feedback and black hole accretion models were calibrated to match the galaxy stellar mass function observed at z $\sim$ 0, providing that this produced galaxies of reasonable sizes (Schaye et al., 2015). The AGN feedback model was calibrated using the observationally derived relationship between black hole and stellar mass (Schaye et al., 2015).

The limits on both the resolution and accuracy of cosmological simulations at the time prevented EAGLE from being able to accurately recreate the cold ISM. The simulations also struggle to determine the efficiency of feedback processes that regulate galaxy growth (Pfeffer et al., 2017). To manage these problems, the subgrid routines are calibrated using the methods previously described.

Despite these limitations, EAGLE is successful in reproducing many observable galaxy characteristics, including the evolution of the stellar masses (Furlong et al., 2015) and

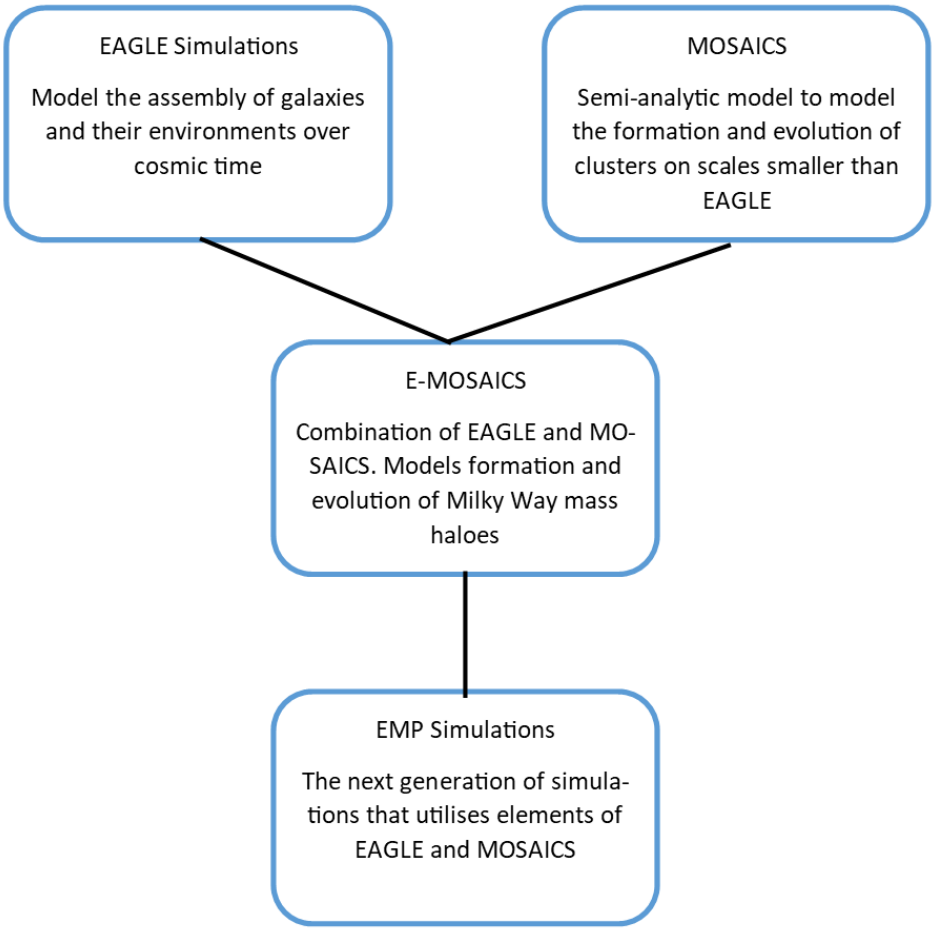

FIGURE 2.1: A 'family tree' illustrating the evolution and combination of models used in this work.

sizes (Furlong et al., 2016) of galaxies, their luminosities and colours (Trayford et al., 2015), their cold gas properties (e.g Crain et al. (2017)), metallicities (De Rossi et al., 2017) and the properties of circumgalactic and intergalactic absorption systems (e.g. Rahmati et al., 2016).

EAGLE was the part of the first generation in a new family of simulations, and the 'family tree' of the simulations used in this work is in Figure 2.1.

2.5 MOSAICS

Whilst cosmological simulations are good for following large scale structure and evolution, the resolution limits mean they cannot model stellar clusters. It was therefore necessary to devise a new model for this purpose.

MOSAICS is a semi-analytic model that is capable of following both the formation and evolution of star clusters (Kruijssen et al., 2019a). It has been used in Kruijssen et al. (2011) and Kruijssen et al. (2012b) to track the formation and evolution of star clusters located in isolated disc galaxies and galaxy mergers. MOSAICS was adapted to include the models described in Kruijssen (2012a) and Reina-Campos & Kruijssen (2017). These additions were necessary to ensure that MOSAICS accurately modelled cluster formation efficiency (CFE) and the maximum cluster mass, as both have an environmental dependence (Pfeffer et al., 2017).

When MOSAICS is run, the main code forms stellar particles which are each assigned a series of star clusters produced by a sub-grid routine (Pfeffer et al., 2017). Due to the association with the stellar particle, all of the clusters have the same coordinates and metallicity as the star particle (Pfeffer et al., 2017). Other birth cluster properties, including CFE, can be calculated based on the properties of nearby stellar particles (Pfeffer et al., 2017).

With the incorporation of MOSAICS into cosmological simulations, it became possible to resolve cluster formation and evolution within large scale simulations.

2.6 E-MOSAICS

E-MOSAICS (MOdelling Star cluster population Assembly In Cosmological Simulations within EAGLE) is a suite of hydrodynamical cosmological galaxy formation simulations run with a combination of EAGLE and MOSAICS (Pfeffer et al., 2018). The simulations are self- consistent and model the formation and evolution of 25 Milky Way mass haloes and their star cluster populations (Pfeffer et al., 2018).

The simulations consist of 25 "zoomed-resimulations", first mentioned in Pfeffer et al. (2018). Each of these is identified as Halo_0XX, with XX representing an integer from 0 to 24. In addition, a full size volume with a box length of 34 Mpc and resolution of 1034^3 is included. This resolution is the same as that used in the initial simulations (Schaye et al., 2015). The cosmological parameters used are outlined in Planck Collaboration et al. (2014a), including $\Omega_0 = 0.307$, $\Omega_b = 0.04825$, and $\Omega_\Lambda = 0.693$.

Galaxy formation follows the prescription set out by EAGLE (Schaye et al., 2015; Crain et al., 2015). The formation and evolution of the star cluster populations is modelled using a modified version of the MOSAICS subgrid model (Pfeffer et al., 2018; Kruijssen et al., 2011).

Star clusters cannot be individually resolved, so they are included as a subgrid population within their host star particle. The formation of these clusters is governed by an environmentally- dependent model based on the fraction of stars formed in bound clusters (Trujillo-Gomez et al., 2021), and the upper truncation mass of the Schechter initial cluster mass function (Reina-Campos & Kruijssen, 2017). Both the bound fraction of stars and the upper truncation mass are dependent on the environment, and both increase with gas pressure (Trujillo-Gomez et al., 2021).

In E-MOSAICS, when a gas particle is converted to a star particle a fraction of the mass is assumed to be in a cluster environment. The fraction is dependent on the pressure of the original gas particle, and is determined by a semi-analytic model that is calibrated using observations of the formation of young massive clusters (YMCs). The fraction of mass is then implemented as a population of globular clusters associated to the newly formed star particle. The masses of the individual globular clusters within the population are randomly drawn so they produce an initial mass function with a slope of -2 that is consistent with Schechter (1976).

After their formation, the clusters lose mass according to the stellar evolution treatment implemented in the EAGLE subgrid models. These follow the model from Wiersma et al. (2009a) and consider four different processes. One process is mass loss through tidal shocks from the interstellar medium. In lower density environments two-body relaxation is the dominant mass loss mechanism (Kruijssen et al., 2011),so this process is also included. For both tidal shocks and two-body relaxation, the instantaneous local tidal field at the position of each particle is used to calculate the mass loss (Trujillo-Gomez et al., 2021). Stellar evolution also contributes to mass loss, and this is calculated with the EAGLE stellar evolution model (Wiersma et al., 2009b).

Clusters can also be impacted by dynamical friction, which is a gravitational interaction between a cluster and objects near by. This interaction causes the cluster to slow down, eventually causing it to spiral towards the galactic centre (Pfeffer et al., 2018). This effect is particularly noticeable for large clusters near the galactic centre (Pfeffer et al., 2018). Dynamical friction is modelled in post- processing and not on-the-fly, as a star particle can host several clusters of different masses, meaning there is not a single timescale over which dynamical friction will occur (Pfeffer et al., 2018). The equation for the timescale for a cluster to spiral to the galactic center is given in Lacey & Cole (1993), and this is

calculated at each snapshot. Clusters are considered to be destroyed if $t_{df} < t_{age}$, where t_{age} is the age of the cluster (Pfeffer et al., 2018).

E-MOSAICS has successfully reproduced several properties associated with globular cluster (GC) populations, including the massive end of the GC mass function (Pfeffer et al., 2018; Hughes et al., 2022), the color-luminosity relation of metal poor GCs (Usher et al., 2018), the GC system mass-halo mass relation (Bastian et al., 2020), the GC radial distribution (Reina-Campos et al., 2022b) and kinematics (Trujillo-Gomez et al., 2021).

2.7 The EMP Simulations

2.7.1 Background

After the success of EAGLE and the incorporation of MOSAICS to create E-MOSAICS, a new generation of simulations has started to form. These simulations incorporate a range of additional physical processes, increasing the accuracy of the original EAGLE simulations. For this work, the Empirically Motivated Physics (EMP) simulations were used, as these are able to model clusters on cosmological scales and include the cold ISM.

The simulations are run on the AREPO code, and the initial conditions are generated using three levels of resolution on the dark matter particles (Reina-Campos et al., 2022a). This means that only the environment within 600 proper kpc of the centre of the target galaxy is simulated in high resolution (Reina-Campos et al., 2022a). The decision was made to use proper kpc as this measures the distance between two points at a given moment in time, and will then evolve with the expansion of the Universe.

2.7.2 Subgrid Models in the EMP Simulations

The spatial resolution in the simulations is set by the gravitational softening, meaning several processes occur on scales much smaller than the spatial resolution (Reina-Campos et al., 2022a). In order to include the physics required to model these processes, subgrid routines are used.

It is not possible to derive all of the parameters used in the subgrid routines from first principles or observations (Reina-Campos et al., 2022a). To decide on values for parameters such as the density threshold n_{th} and the constant star formation efficiency per free-fall time ϵ_{ff}, the initial conditions of halo 4 were evolved using a constant SFE for a range of values for each parameter (Reina-Campos et al., 2022a). The Euclidean

norm of the median errors is calculated to the Baldry et al. (2012) size–mass relation and the Moster et al. (2013) stellar-halo mass relation, and the combination of values that gives the smallest deviation from these results is used (Reina-Campos et al., 2022a). This leads to a SFE per free fall time of 0.2 and a density threshold for star formation of 1 H cm^{-3} (Reina-Campos et al., 2022a).

2.7.2.1 Cluster Formation and Evolution

Cluster formation and evolution is one of the processes conducted via a subgrid routine. As this process requires highly detailed physics, the computational time required is substantial. To reduce both the computational time and cost the model assumes that each stellar particle has a given fraction of mass within gravitationally bound clusters when it forms (Reina-Campos et al., 2022a). As well as reducing the cost, this also enables the investigation into the impact that the galactic environment has on the properties of stellar clusters (Reina-Campos et al., 2022a).

The subgrid model used for this process is an improved version of MOSAICS (Kruijssen et al., 2011; Pfeffer et al., 2018), that has been adapted from that used in E-MOSAICS (Pfeffer et al., 2018; Kruijssen et al., 2019a). The new version of MOSAICS includes five modifications that have been implemented in Reina-Campos et al. (2022a).

The first modification is a new model for the initial cluster mass function, which can cause the minimum cluster mass to increase, and therefore narrow the ICMF (Trujillo-Gomez et al., 2019). This would lead to a reduction in the number of high and low mass clusters, leaving more intermediate mass clusters (Kruijssen et al., 2019a).

The second modification is the inclusion of an environmentally dependent description for the initial half-mass radius (Choksi & Kruijssen, 2021). This model predicts that globular clusters at high redshifts are more compact, making them less likely to be disrupted (Choksi & Kruijssen, 2021).

The third modification is a new cluster disruption description based on N-body simulations (Webb et al., 2019). This describes the impact of tidal shocks on clusters, and predicts a greater effect than those described by analytic derivations. Over a cosmological time scale, this would cause a greater level of cluster disruption than other models.

The fourth modification incorporates a more accurate description of two-body interactions and the disruption they cause (Alexander & Gieles, 2012). This description assumes that clusters are comprised of stars of equal masses, and that the evolution they undergo corresponds to the post core-collapse phase (Reina-Campos et al., 2022a).

These assumptions simplify the dynamics of star clusters by focusing on the stage in their evolution where gravitational interactions and dynamical processes are the most important. This reduces the need for a range of complex physics, minimising the computation time.

The final modification to MOSAICS is the inclusion of the effects of size evolution and how this can influence mass evolution in clusters (Reina-Campos et al., 2022a). Size evolution can impact mass evolution in various ways. For example, if a cluster expands in size, it could incorporate nearby stars, thus increasing in mass. Alternatively, an evolution in the distribution rather than the amount of mass can drive size evolution. For example, dynamical relaxation facilitates the exchange of energy and angular momentum within clusters, leading to more massive stars congregating near the centre of the cluster. This leads to an increase in density and a gravitational contraction.

These modifications to the subgrid model improve the accuracy of the results obtained by the simulations, whilst also ensuring that processes occurring on scales smaller than the resolution are captured appropriately.

2.7.2.2 Star Formation

Star formation also requires a subgrid routine due to the resolution limits (Reina-Campos et al., 2022a). Star formation can be modelled with a constant star formation rate, or an environmentally dependent rate, which is in line with theories such as those outlined in Hennebelle & Chabrier (2011) and Burkhart (2018). In this work, the constant rate was utilised.

For each gas cells in the simulation, the star formation rate is given by:

$$\frac{d\rho_\star}{dt} = \epsilon_{ff} \frac{\rho}{t_{ff}}, \tag{2.3}$$

where ρ is the gas cell density, t_{ff} is the local gas free fall timescale (Reina-Campos et al., 2022a), and ϵ_{ff} is the star formation efficiency per free fall time. We also utilise a Chabrier (2005) IMF to model the distribution of stellar mass.

However, not all gas cells are allowed to undergo star formation. This can only occur when a gas cells has a density over 1 H cm^{-3} and a temperature below 1.5 x 10^{4}K (Reina-Campos et al., 2022a). The high temperature threshold is due to gas in the simulations being unable to cool below 10^{4}K until the first stars have formed (Reina-Campos et al., 2022a). Gas cells must also have overdensities greater than $\delta_{th} = \rho/(\Omega_b \rho_c) = 57.7$, where ρ_c is the critical density in a flat Friedmann universe, and Ω_b is the density parameter

of baryons (Reina-Campos et al., 2022a). By applying this constraint, excessive star formation at high redshifts is avoided (Reina-Campos et al., 2022a).

2.7.3 Feedback Mechanisms and Small Structure Formation

Some evolutionary processes do not require subgrid models. In the EMP simulations, feedback from stellar evolution, AGB (asymptotic giant branch) winds, and supernovae are included without the need for subgrids, allowing the chemical properties of the ISM and the star particles to be tracked over time (Reina-Campos et al., 2022a).

As previously mentioned, the resolution of the EMP simulations is limited, and as such it does not allow for individual star and planet forming systems to be resolved. Star particles can be resolved and provide a good proxy for stars that is used in this work. For this research, the main variables used were the cluster mass, field mass, lookback time, and metallicity of each particle.

As mentioned above, previous attempts at simulating cluster formation and evolution from the Big Bang to the present have lacked the accuracy desired, primarily due to the cold ISM being left out or inaccurately modelled. The EMP simulations include a more accurate description of the cold ISM, producing more accurate models that will demonstrate if and how the cold ISM impacts the formation of clusters and stars. The models will also provide evidence of any impact the cold ISM has on the lifetimes of proto-planetary disks.

2.7.4 Initial Conditions

The EMP simulations were designed to study galaxies with masses and luminosities similar to the Milky Way (Reina-Campos et al., 2022a). The aim was to track the formation and evolution of the stellar clusters in these galaxies when the cold, clumpy ISM was modelled (Reina-Campos et al., 2022a). Therefore, the EMP simulations consist of a suite of zoom-in simulations of Milky Way mass galaxies. The initial conditions of these galaxies are the same as those used in the volume limited sample from E-MOSAICS (Pfeffer et al., 2018; Kruijssen et al., 2019a).

This sample was drawn from the EAGLE Recal-L025N0752 DM-only periodic volume (Schaye et al., 2015) after applying a cut of $11.85 < \log_{10}(\mathrm{M}_{200}/M_{\odot}) < 12.48$ at the present day to the halo mass (Reina-Campos et al., 2022a). The region around each halo was identified as a region of interest (ROI) in the dark matter only periodic volume (Reina-Campos et al., 2022a). As the original volume is periodic, the simulated universe is the same in all directions, and there is no need to inspect a larger region.

After finding the ROI, the zoomed initial conditions were created at z = 127 for each halo (Reina-Campos et al., 2022a). This is done using the Gaussian code PANPHASIA (Jenkins, 2013) and the second order Lagrangian perturbation theory method (Jenkins, 2010). The first step of this method is grid initialisation, which uses a regularly spaced grid to represent the initial density field (Jenkins, 2010). Each grid cell is of an equal size and represents a particle. The first-order linear displacement field is then calculated based on the assumption that the density perturbations will be small (Jenkins, 2010). This then allows the density field to be evolved over time. The linear displacement field is then used to calculate the first order displacements of the particles from their starting points (Jenkins, 2013). The gradient of the first order displacement field is then calculated to give the velocity field, which gives the initial velocities of the particles due to the gravitational forces of other particles (Jenkins, 2013). Second order corrections are then calculated and applied to the displacement field in order to account for non-linear effects, such as particles crossing each other's trajectories. Second order corrections are also calculated for the displacements to obtain accurate final displacements for each particle (Jenkins, 2010). Based on these final displacements, the positions of the particles can now be calculated, giving the initial conditions of the particles for the simulations.

The cosmological parameters used are the same as those provided by Planck (Planck Collaboration et al., 2014b) and are outlined in Table 2.1.

TABLE 2.1: Cosmological parameters as determined by Planck

Parameter	Symbol	Value
Matter Density Parameter	Ω_m	0.307
Baryon Density Parameter	Ω_b	0.048
Dark Energy Density Parameter	Ω_Λ	0.693
Density Fluctuation Amplitude	σ_8	0.829
Hubble Constant	H_0	100h km s^{-1} Mpc^{-1}
Hubble Constant	h	0.677

2.8 Statistical Tests

Due to the scales covered in this work and the amount of data utilised, it is not possible to track and record the individual evolutionary histories of each star particle and compare this to solar type particles. Therefore, statistical tests were employed to provide insight into the similarities and differences between particle populations. Specifically, the two-sample Anderson-Darling (AD) and Kolmogorov-Smirnov (KS) tests were used. These tests are non-parametric statistical methods for comparing two distributions. Despite their common goal of assessing whether two sets of data come from the same distribution, these tests differ in their approaches. Due to these differences, the Kolmogorov- Smirnov

test is generally better at identifying similarities in the main bodies of distributions, whereas the Anderson- Darling test is better at identifying similarities in the tails. Therefore, we utilise both tests to ensure that distributions are truly similar.

The Kolmogorov-Smirnov test uses a null hypothesis that both samples are drawn from the same underlying distribution. The first step is to calculate the KS statistic, which is the maximum absolute difference between the 2 empirical cumulative distribution functions of the samples. This distance is always taken to be the largest vertical distance between the two functions, even if there is a larger horizontal difference. The KS statistic is calculated via:

$$D = max|F_1(x) - F_2(x)|, \tag{2.4}$$

where D is the KS statistic, and $F_1(x)$ and $F_2(x)$ are the empirical cumulative distribution function of the two samples.

Once the KS statistic has been calculated, it can either be compared to a critical value from the Kolmogorov-Smirnov distribution table, or a p-value can be manually calculated using the equation:

$$p\text{-}value = 2\left(1 - \Phi D\sqrt{\frac{n_1 n_2}{n_1 + n_2}}\right) \tag{2.5}$$

where D is the KS statistic, Φ is the cumulative distribution function of the standard normal distribution, and n_1 and n_2 are the sizes of the two samples. For this work, a Python package is used to calculate the p-value. Whichever method is used, both the critical value from the table and a calculated p-value provide the probability of obtaining a KS statistic like the one observed if the null hypothesis is true.

The Anderson-Darling test uses the null hypothesis that both samples are drawn from the same distribution. For each sample used in the Anderson-Darling test, the difference between the empirical distribution function (EDF) and the cumulative distribution function of the hypothesised distribution is computed. The sum of the weighted square differences between the observed cumulative density functions (CDFs) and the expected CDF is calculated, and if this is greater than the critical value in the table used by the Python package, the null hypothesis can be rejected.

Both tests utilise the differences in the shape of the distribution when compared to a reference distribution to perform their analysis. However, the method used by the Anderson-Darling test puts additional emphasis on the tails of the distribution compared to the Kolmogorov-Smirnov test. This is because the Anderson-Darling test utilises a

weighting factor, with data points at the tails being assigned a higher weight. For this reason, both the Anderson-Darling and Kolmogorov-Smirnov tests are used to analyse the results.

For this work, the $\log_{10}$ of the p-values produced by both tests is taken. In Python, there are no limits on the p-values of the KS test, but the AD test is subjected to a cap of 0.25 and a floor of 0.001. This limits the $\log_{10}$ p-values to between -3.0 and -0.6021. To determine a reasonable cut off for the statistical tests, both 2, and 3 sigma values were trialled. As the number of results included using both of these values did not vary significantly, a 2σ cut off of -1.30 was chosen. This is applied to both the Kolmogorov-Smirnov and Anderson-Darling tests. Any values below this are discounted as statistically insignificant.

We also highlight results that may have a limited validity due to a lack of particles in one or both groups. As both KS and AD tests compare distributions, a distribution consisting of a small number of particles can easily be seen to mimic another distribution due to a lack of points available for comparison. This could result in a p value that suggests a greater level of similarity than is actually present. Ideally, only groups of a similar size would be compared to each other. However, this is not always possible. It is therefore necessary to decide when a group would be considered to be of limited statistical value. For this work, a group was considered to be of limited value when it contained less than 100 particles. This is because most groups contained over 1000 particles, meaning any groups with less than 100 particles were likely to be too small to obtain accurate statistical values from the tests. Therefore, any results where one or both groups contain less than 100 particles have been considered limited, and, where these values are presented, they have been italicised.

Chapter 3

Calculating the Variables Used for Analysis

This work has been carried out on 2 scales. Initially the work was done on particles in the central Milky Way type galaxy and the associated halo. The analysis was then repeated on just the Milky Way type galaxy, removing any particles that were only gravitationally bound to the halo but not included in the galaxy. The process to calculate the variables was very similar in both cases, so in this chapter I will explain the processes and highlight any differences.

As there are already some well known trends that link certain characteristics of stars, such as metallicity and halocentric radius (Doner et al., 2023), this work has examined these trends to ensure they are reproduced, as well as examining other variables to determine if there are other trends between characteristics.

3.1 Format of the Data and Identifying Haloes and Galaxies

16 runs from the current EMP Pathfinder data set are used in this work. In the outputs, each run is referred to as Halo X, with X being an integer from 1 to 24. To avoid confusion between simulation runs and astrophysical haloes, we will adopt the name 'Run X' from here on in. Some runs have been discarded as they were terminated before the simulation was complete. Within each run is a single Milky Way type galaxy which may also have satellites. As this research is focused on solar type stars in Milky Way type galaxies, any satellites have been discarded.

For this work, the Friends-of-Friends (FoF) algorithm (Davis et al., 1985) was used to identify structures that have formed due to the collapse of dark matter. The algorithm identifies dark matter haloes before identifying gas and star particles and associating these to the nearest dark matter particle. As each dark matter particle is associated to a dark matter halo, the gas and star particles are then automatically associated with the appropriate halo (Reina-Campos et al., 2022a).

The SUBFIND algorithm Springel et al. (2001) is then used to identify substructures within the haloes that have become gravitationally bound (Reina-Campos et al., 2022a). These are the galaxies within the halo. The central, Milky Way type galaxy is assigned a FoF ID, which is used in a mask to remove any particles not bound to both the FoF halo and the central galaxy.

To identify particles that were bound exclusively to the the central galaxy, it was necessary to first calculate the scale height of each Milky Way type galaxy. This was done using pynbody, which is a Python package designed to calculate a range of physical properties for N-body simulations (Pontzen et al., 2013). In this work, pynbody was used to rotate the Milky Way type galaxy in each run so that z = 0 defined the midplane. The absolute value of the z coordinate of each particle was plotted in a histogram, and the point at which the number of particles had decreased by a factor of e was determined, providing the scale height of the galaxy. Any particles with an absolute z value above the scale height were removed, leaving only particles considered to be part of the Milky Way type galaxy.

3.2 Classifying the Star Particles

As this work aims to identify the impact that early environments has on the formation and evolution of stars similar to the Sun, it was necessary to define the properties of a solar type particle. The simulation runs calculate the age and metallicity of each particle, which can be used to define a solar type particle. To determine if there is any obvious structure in the age-metallicity plane of each halo, contour plots were created. An example of this can be found in 3.1. As these plots did not show areas of unusually high or low stellar density, there were no obvious cuts to be made for either age or metallicity. Therefore, we selected ranges that are in line with currently accepted values. Star particles with ages similar to the Sun were defined as having ages between 4.3-4.7 Gyrs (Bonanno et al., 2002). The metallicity range used to classify a star particle as having solar metallicty was -0.2 dex $\geq$ [Fe/H] $\leq$ 0.2 dex.

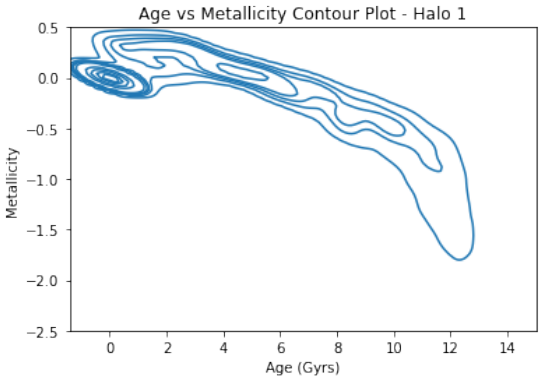

FIGURE 3.1: Contour plot in the age-metallicity plane for Halo 1. As with the plots for other haloes, this plot showed no areas with unusual densities.

TABLE 3.1: Details of the metallicity and age cuts used to classify star particles.

Category	Cut Applied
Young	Age $<$ 4.3 Gyrs
Solar Age	4.3 Gyrs $\leq$ Age $\leq$ 4.7 Gyrs
Old	Age $>$ 4.7 Gyrs
Metal Rich	[Fe/H] $>$ 0.2
Solar Metallicity	-0.2 $\leq$ [Fe/H] $\leq$ 0.2
Metal Poor	[Fe/H] $<$ -0.2

These ranges ensures that the majority of runs contained enough solar type star particles for statistically robust comparisons to be made, without including so many star particles that the solar group lacked a level of homogeneity. The final cuts used to classify the particles can be found in Table 3.1.

These cuts created 3 metallicity groups (solar metallicity, metal rich, and metal poor), and 3 age groups (solar age, old, and young), giving 9 age-metallicity combinations that were utilised for this work. From these cuts, we define a solar type particle as a particle which has a solar age and metallicity.

Classifying the particles in this way allows the properties of solar type particles to be compared with the properties of particles in other groups. Specifically, we can establish i) if particles with solar age and metallicity experience significantly different evolution histories compared to other particles, and ii) if there are any trends in evolutionary histories that could be linked to either age or metallicity.

3.3 Halocentric Radius Distribution of Star Particles at z = 0

As there are well known trends linking halocentric radius, age, and metallicity (e.g Shields, 1974; Doner et al., 2023). For example, research has repeatedly demonstrated a negative metallicity gradient with respect to halocentric radius (e.g Anders, F. et al., 2017; Doner et al., 2023). This means that the present day halocentric radius of a particle can encode information about its evolutionary history.

Our simulations output the 3D spatial coordinates for each particle, which are given relative to the centre of potential of the central galaxy. The halocentric radius (GCR) at z = 0 for each particle was calculated using the x,y,z coordinates, as shown in 3.1.

$$GCR = \sqrt{x^2 + y^2 + z^2}, \tag{3.1}$$

The distance distribution of the star particles from each of the 9 categories is used in both a Kolmogorov-Smirnov and Anderson-Darling test to establish how likely it is that each group is drawn from the same distribution as the solar particles.

3.4 Calculating Mass-Weighted Time in Clusters

Cluster environments are considered to be hostile to planets. For example, photoevaporation from nearby stars will heat protoplanetary disks and prevent the formation of planets. The range of dynamic events occurring in cluster environments is known to have an impact on the evolutionary paths, and therefore the present day properties, of stars (Maraboli et al., 2023). This makes it an important metric to consider in this work.

Therefore, we aim to compare the typical amount of time star particles from different groups spend in a cluster environment. However, the initial stellar mass in each particle varies at the time of formation, and the distribution of this mass (between field and clusters) varies between individual particles and groups. In order to compare star particles with different mass distributions, we create a new metric called the mass-weighted time (MWT). MWT is calculated using Equation 3.2.

$$MWT(x) = \frac{M_{cl}(x) + M_{cl}(x+1)}{2}\Delta t, \tag{3.2}$$

where M_{cl}(x) is the fraction of mass in clusters for snapshot x, M_{cl}(x+1) is the fraction of mass in clusters for snapshot (x+1), and Δt is the time interval between snapshot(x) and snapshot(x+1). This is the average fractional stellar mass at a given time step multiplied by the time step.

The EMP Pathfinder simulations calculate the amount of stellar mass in a gravitationally bound cluster and the stellar mass in the field for each star particle at its formation and every subsequent snapshot. The fraction of mass in a cluster is therefore found by dividing the former by the sum of the former and latter.

The time interval, Δt is the difference in lookback times between snapshot(x) and snapshot(x+1). As the simulations output at equal redshift intervals, the snapshots do not occur at equal time intervals. Therefore, Δt needed to be calculated separately for each value of x. When a star particle does not form in a snapshot, the lookback time of snapshot(x) is replaced by the formation time of the particle, and the lookback time of snapshot(x+1) is replaced by the lookback time of the first snapshot taken after the particle is formed.

This process gave a fraction of mass contained in clusters and a time interval that are multiplied to give the mass-weighted time in clusters for a given snapshot. This is repeated for all subsequent snapshots that the star particle appears in, and the sum of these values gives the MWT value for the particle. This is then repeated for all particles, and the values are plotted in a cumulative frequency plot.

The distributions of MWT for each of the groups are compared to that of the solar group in the same halo. We then compare the solar groups from different haloes to each other. Both of these processes are carried out with both KS and AD tests.

3.5 Mass in a Cluster Environment Since Formation Time

A key part of our analysis involves understanding the early evolutionary histories of star particles. This requires sampling the properties of the star particles on a common time scale. As previously mentioned, the simulations do not use regular time intervals, instead opting for regular redshift intervals, meaning the variables associated with each particle are only regularly sampled in redshift. Coupled with this, particles are not limited as to when they can form, meaning there is no guarantee that a particle will form in a snapshot. It is therefore necessary to interpolate the properties, specifically the fraction of mass in a cluster environment, of each particle on to a shared time grid.

To carry out the interpolation, it was first necessary to establish how much time had passed since the formation of each star particle. This was done using the cumulative summation of the Δt values calculated during the investigation into the mass-weighted time. These values were entered into an array, with each row representing a star particle within the group. A second array held the fraction of mass in a cluster environment at these times. Interpolating both arrays provided the fraction of mass in a cluster environment for each particle since it had formed. The interpolation was carried out 3 times: once for the entire simulation run with time steps of 0.5 Gyrs, once for the first gigayear of the simulation with a time step of 0.1 Gyrs, and finally for the first 100 Myrs with a time step of 10 Myrs.

The interpolated data allows for the comparison of particles from different groups. However, each group can contain tens of thousands of star particles, so to compare the trends between groups, we take the mean for each group. This was done for the mass in a cluster environment for each of the 9 groups in each halo.

3.6 Expectation Values for the Amount of Time in a Cluster Environment

Works such as Pfalzner, S. (2013) and Gonzalez (2005) outline the impact that the environment can have on the properties of a star. Stars that spend prolonged periods of time in hostile environments are therefore likely to have been impacted to a greater extent. Because of this, it is important to determine if particles from different groups spend different amounts of time in a cluster environment. To compare the particles, we calculate expectation values for each group to quantify the amount of time spent in a cluster environment.

The general form of the expectation value is given by 3.3:

$$E(X) = \sum xP(x) \tag{3.3}$$

where x is the values of variable X, and P(x) is the probability.

As we are interested in the expected amount of time in a cluster environment, we replace x and P(x) with the interpolated time and the interpolated fraction of mass in a cluster environment. This gives the equation:

$$E(X) = \sum T_i CM_{frac} \tag{3.4}$$

where T_i is the interpolated time, and CM_{frac} is the interpolated fraction of mass in a cluster environment.

To calculate the expectation values for the amount of time spent in a cluster environment, the interpolated fraction of mass in a cluster environment is utilised, along with the interpolated time for the first 100 Myrs. For each particle, the difference in the fraction of mass in a cluster environment is calculated for each time step. These values are then multiplied by the 10 Myrs time step, and the sum of the values for each particle is taken. The mean for each of the 9 groups is then taken and converted into Myrs. This gives the expectation value for the amount of time stellar mass will spend in a cluster environment during the first 100 Myrs since formation.

3.7 Expected Number of Supernovae in the First 100 Myrs Since Formation

When stars go supernova, they release a large amount of energy into their local environment. This energy can seriously impact any nearby stars (Close & Pittard, 2017), and can destroy their protoplanetary disks. As any such interactions could impact how stars evolve, it was important to determine how many supernova events a given star was likely to undergo. As supernovae are significantly more common at early times due to the large number of O and B type stars, this analysis is limited to the first 100 Myrs since the formation of each particle.

By using the same interpolation method described previously, the amount of mass ($M_\odot$) in a cluster environment at 10 Myr intervals in the first 100 Myrs since a particle's formation is determined.

Using the data from Figure 43 of the Starburst99 model (Leitherer et al., 1999), the average rate of supernovae for a 10^6 $M_\odot$ cluster is 10^{-3} per year. Therefore, the number of supernovae experienced in a 10 Myr time step can be calculated via:

$$SN = 10^{-3} \times \frac{M[M_\odot]}{10^6} \times 10^7 \tag{3.5}$$

However, as the EMP simulations do not resolve individual clusters, these results must be discussed in terms of particles. Therefore, these values are multiplied by the fraction of mass in a cluster environment at each time step to give the total number of supernovae expected in the time step for each particle. The total number of supernovae over the first 100 Myrs is then taken for each particle. This allows the solar group to be compared to

the 8 other groups to determine if solar particles are likely to experience a significantly different number of supernova in their early lives. As with previous statistical analysis, this is done using KS and AD tests.

Chapter 4

Results from the Halo Analysis

It is possible to analyse haloes in isolation, as well as comparing them to each other. By initially analysing the haloes individually, it is possible to determine if solar type particles show a unique evolutionary history. If this were found to be true, it may suggest that the Sun evolved in a way particularly supportive of life. If strong overlaps between the evolutionary histories of solar type particles and other groups are found, it would suggest that particles with different ages and metallicities may also share a similar evolutionary history, suggesting that they may also be able to support life. This would potentially suggest that none of the variables analysed have a major impact on the formation of planets or future habitability, and that the conditions leading to the formation of the solar system are more random, or occur on a more local scale.

As this work is focused on establishing if solar type particles have different evolutionary histories to other particles, only the solar group was compared between haloes. Comparing the solar type particles from different haloes allows trends between haloes to be identified.

In both cases, the halo-centric radius, mass-weighted time, the expected number of supernovae experienced in the first 100 Myrs since formation, and the amount of time spent in a cluster environment in the first 100 Myrs were analysed using both Kolmogorov-Smirnov and Anderson-Darling tests (see Section 2.8). As there are known trends between stellar density and radius, and metallicity and radius, analysing the halo-centric radius can indicate the likely metallicity of a star and also give some indication as to its proximity to supernovae events. As discussed in Section 1.3.4, supernovae are crucial for the synthesis of the heavy elements used in planet formation, but are also energetic enough to destroy planets and their atmospheres (e.g. Melott et al., 2017; Portegies Zwart, S. et al., 2018). As cluster environments have higher stellar densities, and more

stars going supernova, remaining in clusters for a prolonged period may alter the habitability of planets.

As these variables are all linked to the overall habitability of a star and its planets, all these variables are analysed. This section describes the processes used to establish the accuracy of the simulations, analyse individual haloes, and compare solar groups across the haloes.

As the haloes are not identical, Table 4.1 provides a summary of their properties. This has been adapted from Reina-Campos et al. (2022a), where further details can be found.

TABLE 4.1: The properties of each halo. The columns give the halo number, halo mass, stellar mass, stellar half-mass radius, and recent star formation rates.

Halo	M_{200} ($10^{12} M_\odot$)	M_* ($10^{10} M_\odot$)	$R_{h,*}$ (kpc)	SFR ($M_\odot$ yr^{-1})
1	1.19	2.33	1.9	6.13
2	1.85	2.77	1.1	3.94
3	1.30	2.40	2.1	3.80
4	0.97	1.02	2.4	0.89
5	1.51	2.77	3.5	4.50
6	0.81	0.69	3.8	0.80
10	2.25	4.58	1.6	8.03
11	1.38	1.26	3.8	2.02
12	2.03	6.05	2.0	12.65
13	2.22	6.23	1.4	15.89
14	2.11	3.90	5.0	13.66
15	1.31	1.04	5.0	2.49
20	0.95	0.88	6.9	2.81
22	1.57	6.43	3.4	11.37
23	1.34	4.21	0.7	5.38
24	1.17	2.97	2.6	7.75

4.1 Using Metalliciy Gradients to Identify Major Mergers

As this work is focussed on Milky Way type haloes, it is important that the metallicity gradients of the haloes mimic that of the Milky Way. There is an established link between the Galactocentric radius and metallicity, which is illustrated by a negative metallicity gradient (Doner et al., 2023). As the centre of each Milky-Way type galaxy was also the centre of each halo, the metallicity gradients in each halo and galaxy were plotted to confirm that they are similar to that of the Milky Way. Haloes found not to have a metallicity gradient similar to the Milky Way will be further analysed to determine the cause of this deviation.

To identify any major mergers, the mean distance from the halo centre of each age-metallicity group in each halo was taken. As the haloes are not the same size, these

values were then divided by the maximum halo radius, giving a fractional distance that allows for direct comparisons between haloes. The results are found in Table 4.2.

TABLE 4.2: Mean Percentage values for the fractional radius from the halo centre of each group in each halo. Results in italics represent a group with less than 100 particles, and results in bold indicate a merger occurred in the halo.

Halo Number	Mean Distance from Halo Centre as a Percentage of Maximum Halo Radius							
	Old Rich	Solar Rich	Young Rich	Old Solar	Solar	Young Solar	Old Poor	Solar Poor	Young Poor
1	0.224	0.136	0.116	0.545	0.618	1.067	2.780	n/a	*3.501*
2	0.114	0.101	0.075	0.289	*0.281*	0.377	3.767	*25.631*	5.025
3	1.182	0.650	0.531	2.296	1.828	2.734	3.122	n/a	n/a
4	*0.251*	*0.273*	0.145	0.442	0.394	0.632	3.139	*25.605*	10.508
5	2.713	1.949	1.560	4.147	4.346	5.991	7.284	n/a	n/a
6	*0.348*	*0.252*	*0.560*	0.736	1.231	1.394	3.188	*2.868*	4.145
10	0.086	0.054	0.047	0.254	0.209	0.533	2.457	0.744	1.397
11	n/a	n/a	*0.240*	0.324	*0.278*	0.352	3.392	0.996	0.959
12	0.106	0.088	0.077	0.332	0.318	0.852	3.001	1.905	2.929
13	0.090	0.068	0.057	0.316	0.484	0.476	1.670	1.550	1.571
14	*0.159*	n/a	0.052	0.373	0.653	0.713	2.684	4.396	2.084
15	n/a	*0.864*	0.999	4.404	3.784	2.859	5.775	4.624	5.421
20	n/a	n/a	n/a	0.442	0.356	0.416	3.320	2.774	3.257
22	**7.421**	**7.668**	**4.427**	**1.427**	**1.169**	**0.453**	**1.328**	**1.376**	**1.565**
23	1.089	0.362	0.146	2.793	1.937	3.966	4.122	n/a	n/a
24	*0.159*	*0.152*	0.104	0.470	0.609	1.037	3.836	9.304	5.848

No errors are included in Table 4.2 as they are all of the order 10^{-3} or 10^{-4}, meaning they are not significant in the comparison of the tabulated value.

It was noted that the values listed for all of the groups in halo 5 were around ten times larger than those for the other haloes. In addition to this, some of the values in other haloes (e.g. solar age metal poor particles in halo 4) were also significantly higher than most of the others. As the mean values in all of the haloes could be skewed by some outliers, the median values were also calculated so they could be compared to the mean values.

The use of median values reduced the values seen in halo 5, suggesting that the mean values in this halo are skewed by outliers. The median values of halo 11 showed that young metal rich particles had a higher median value than both old solar and solar particles. However, the young metal rich group has less than 100 particles, which may cause this discrepancy. Halo 15 also shows some discrepancies, but these are only apparent when considering the median values to 3 decimal places.

When the mean values are considered, all haloes show a negative metallicity gradient, with the exception of halo 22. This supports the results of previous studies of galactic metal gradients (e.g. Doner et al., 2023). Halo 22 generally has solar metallicity particles with the smallest mean radius, followed by metal poor particles and then metal rich particles.

Three examples of distance distributions can be found in Figures 4.1, 4.2, and 4.3.

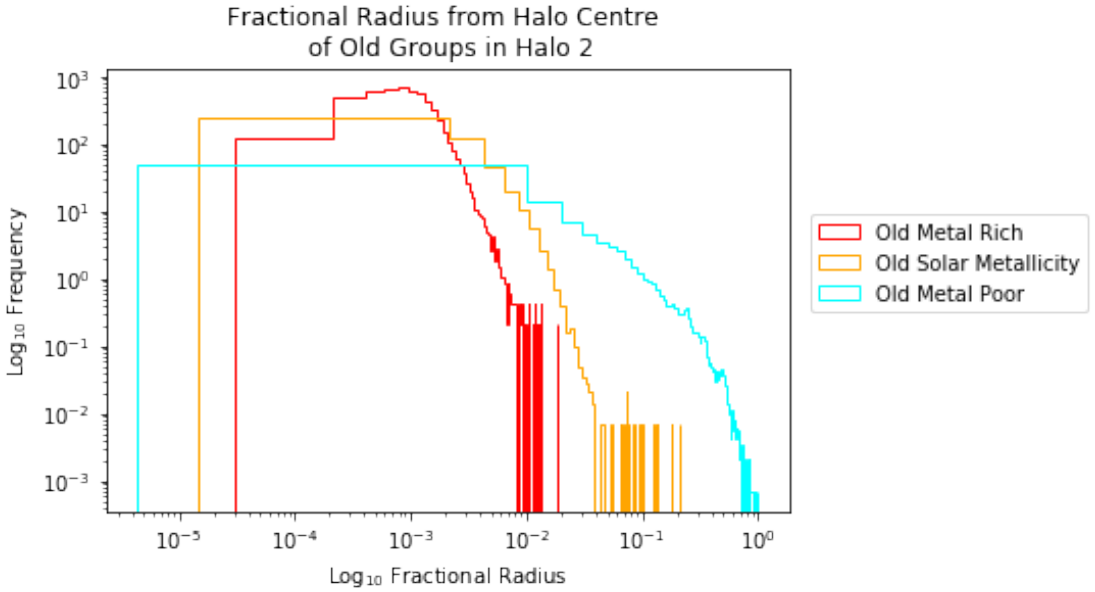

FIGURE 4.1: Fractional distance distribution of old particles in halo 2

TABLE 4.3: Median percentage values for the fractional radius from the halo centre of each group in each halo. Results in italics represent a group with less than 100 particles, and results in bold indicate a merger occurred in the halo.

Halo Number	Median Halo-Centric Radius as a Percentage of Maximum Halo Radius							
	Old Rich	Solar Rich	Young Rich	Old Solar	Solar	Young Solar	Old Poor	Solar Poor	Young Poor
1	0.107	0.090	0.091	0.320	0.473	0.862	0.882	n/a	*3.699*
2	0.099	0.090	0.068	0.196	*0.272*	0.296	1.008	*26.194*	5.068
3	0.637	0.277	0.181	1.181	0.852	1.434	1.600	n/a	n/a
4	*0.172*	*0.257*	0.108	0.329	0.272	0.384	1.143	*28.825*	3.358
5	1.229	0.917	0.516	2.123	2.314	3.476	4.441	n/a	n/a
6	*0.282*	*0.269*	*0.382*	0.549	1.020	1.150	1.258	*2.963*	4.134
10	0.062	0.048	0.044	0.118	0.172	0.456	1.042	0.663	1.334
11	n/a	n/a	*0.264*	0.234	*0.199*	0.265	0.986	0.467	0.719
12	0.091	0.079	0.070	0.223	0.225	0.641	0.975	1.360	2.725
13	0.074	0.060	0.052	0.227	0.426	0.411	0.488	1.498	1.532
14	*0.121*	n/a	0.043	0.219	0.388	0.452	0.975	1.405	1.621
15	n/a	*0.864*	0.471	1.808	1.627	1.068	2.440	1.807	2.178
20	n/a	n/a	n/a	0.350	0.286	0.317	1.529	1.111	2.555
22	**5.317**	**5.339**	**3.508**	**0.296**	**0.356**	**0.214**	**1.124**	**1.290**	**1.527**
23	0.208	0.081	0.044	0.655	0.455	1.164	1.262	n/a	n/a
24	*0.142*	*0.155*	0.083	0.316	0.412	0.645	1.622	5.912	3.727

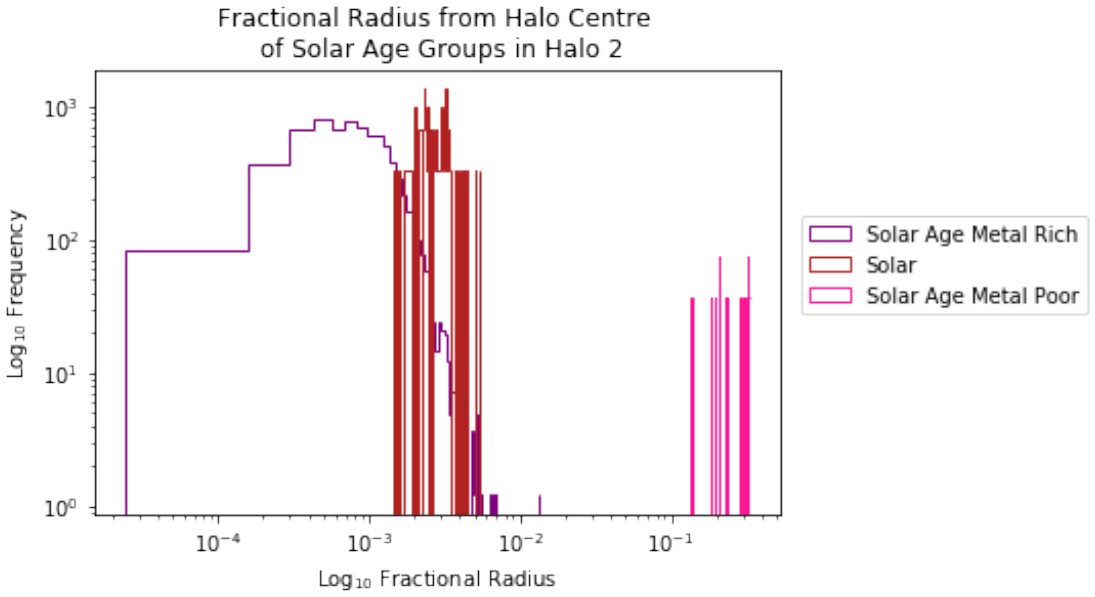

FIGURE 4.2: Fractional distance distribution of solar age particles in halo 2

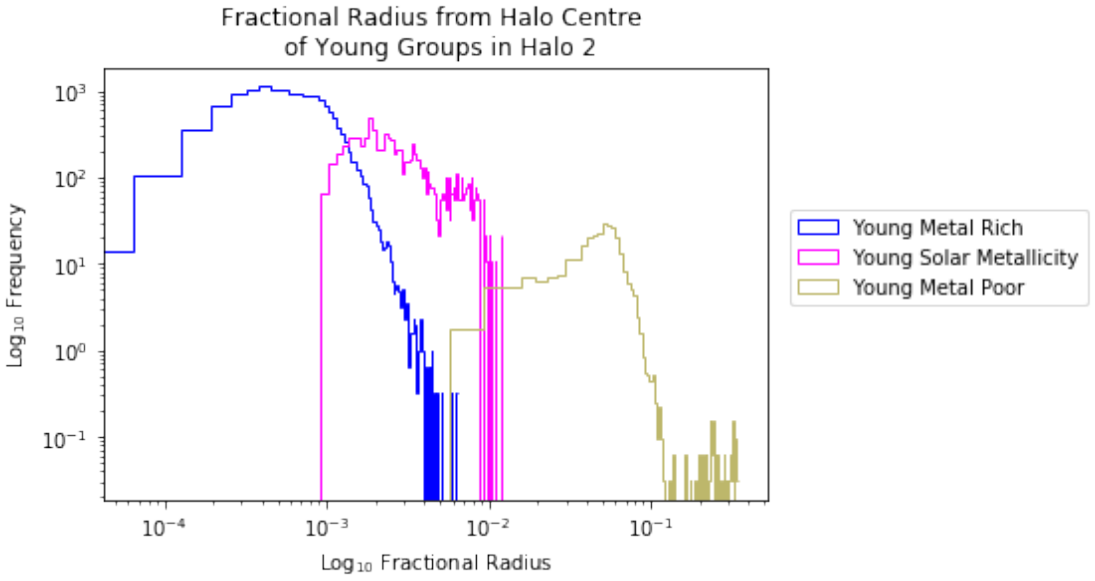

FIGURE 4.3: Fractional distance distribution of young particles in halo 2

Each of these figures shows the distance distributions of particles of the same age from the same halo. In each case, metal rich particles are found closest to the centre of the halo, followed by solar metallicity particles, and then metal poor particles. However, there are often overlaps between the distributions, suggesting that the different metallicity populations are not strictly separated, and that the trend seen is a bulk property of the populations.

4.1.1 Merger in Halo 22

As the metallicity gradient of Milky Way type galaxies is reasonably well understood, the result obtained from halo 22 was unexpected. A change in the metallicity gradient of a halo can be indicative of a second population joining the native population. If this second population joined simultaneously (e.g. via a merger) they would be likely to occupy the same region of space, altering the properties of the halo, such as metallicity and particle velocity. This has not occurred in the Milky Way for around 3 Gyrs (Donlon et al., 2024). To identify if this was the case, or if this was indicative of a problem with the simulation, the change in the metallicity gradient was investigated further.

Initially, the metallicity gradient within the halo was plotted, as shown in Figure 4.4. This highlighted a negative gradient within the first 10 kpc, before flattening off and then increasing at around 15 kpc. From this, it was established that any particles that had merged with the halo were likely to be found 10-15 kpc from the centre of the halo.

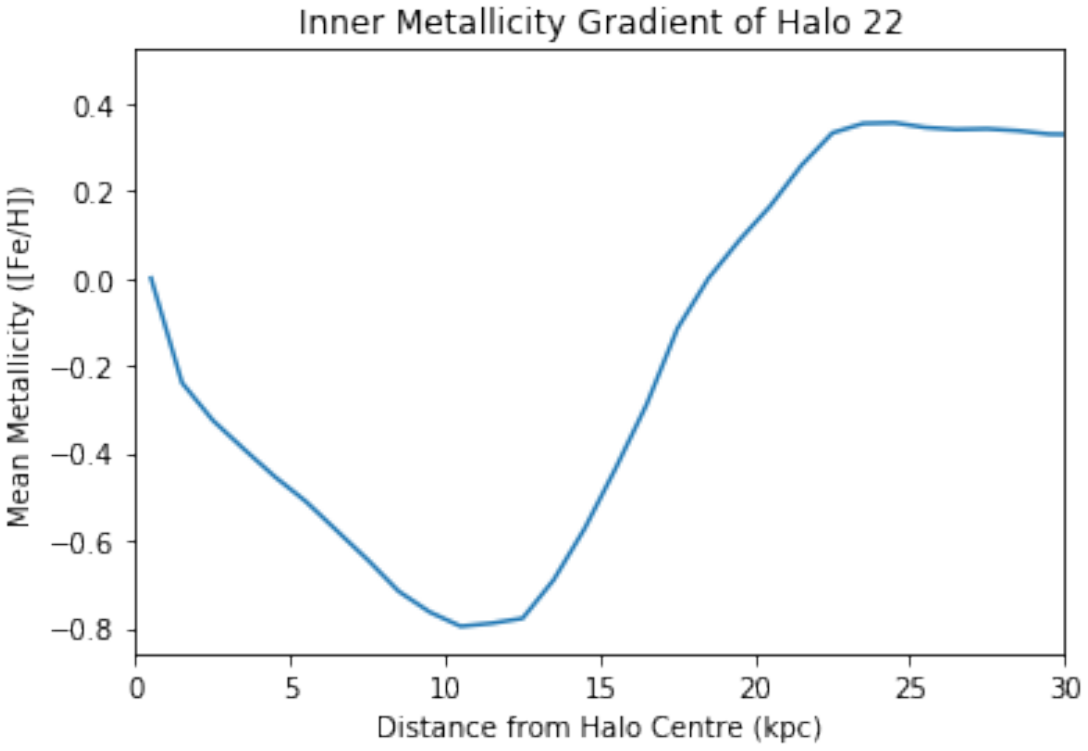

FIGURE 4.4: Metallicity gradient for the inner 30 kpc of Halo 22.

Particles that had merged into the halo were likely to have different velocities when compared to local particles from the native population, and were also likely to be found grouped together. In order to identify any particles that fit these criteria, the velocity of the particles was calculated in cylindrical coordinates. Figure 4.5 shows a plot of v_ϕ vs v_r, with the particles coloured based on their distance from the centre of the halo. This

highlighted a group of particles with velocities usually found in the native population at smaller distances, suggesting that this group of particles had joined the halo in a major merger event. By contrast, Figure 4.6 shows no populations with significantly different velocities, suggesting that there has not been a major merger. As the Milky Way has not undergone any major mergers (Kruijssen et al., 2020), halo 22 must have a different assembly history to the Milky Way. The focus of this work is on Milky Way analogues, so it would be possible to remove halo 22 from this analysis. However, based on measurements from Hubble, it is predicted that the Milky Way and Andromeda will collide, meaning a major merger will occur in the future (Sohn et al., 2012; van der Marel et al., 2012). Therefore, halo 22 represents a potential outcome of this merger, and is still included in this work.

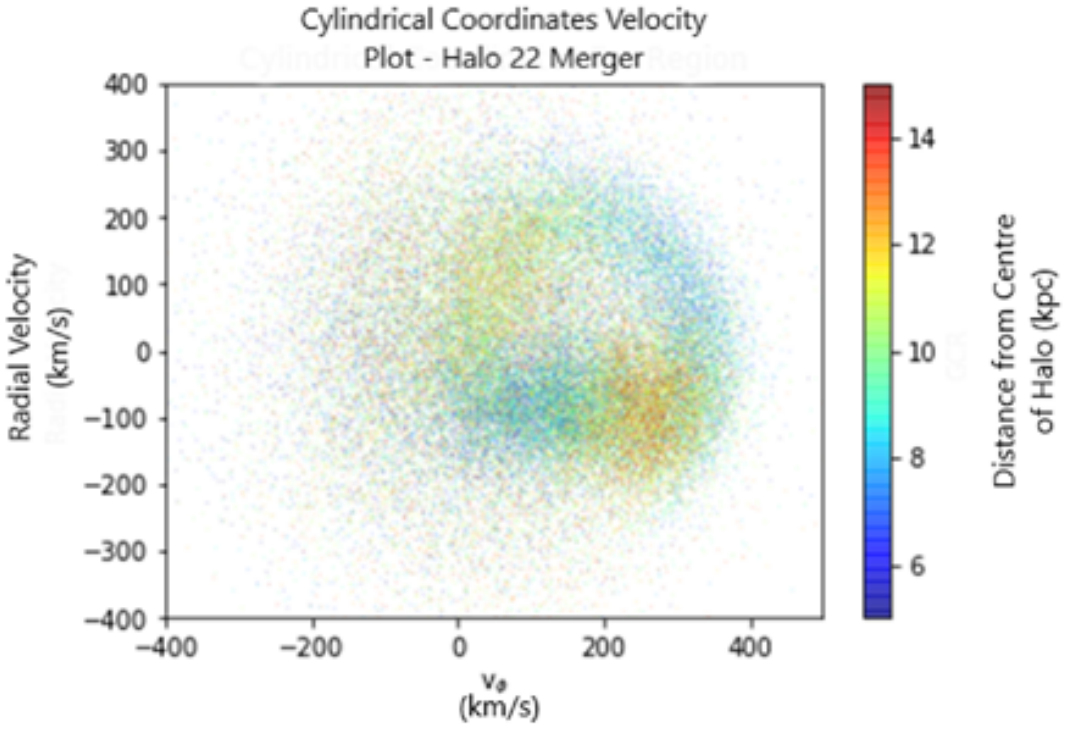

FIGURE 4.5: Velocity plot for particles 10-15kpc from the centre of Halo 22. The collection of red points in the bottom right show particles gained through a merger.

It is also important to note that this merger has introduced another population into halo 22. When KS and AD tests are used, it is assumed that one population is being compared to another, but due to the merger, this is not true. This means that the results of statistical tests involving halo 22 are compromised. Therefore, the results obtained from halo 22 are presented in bold type.

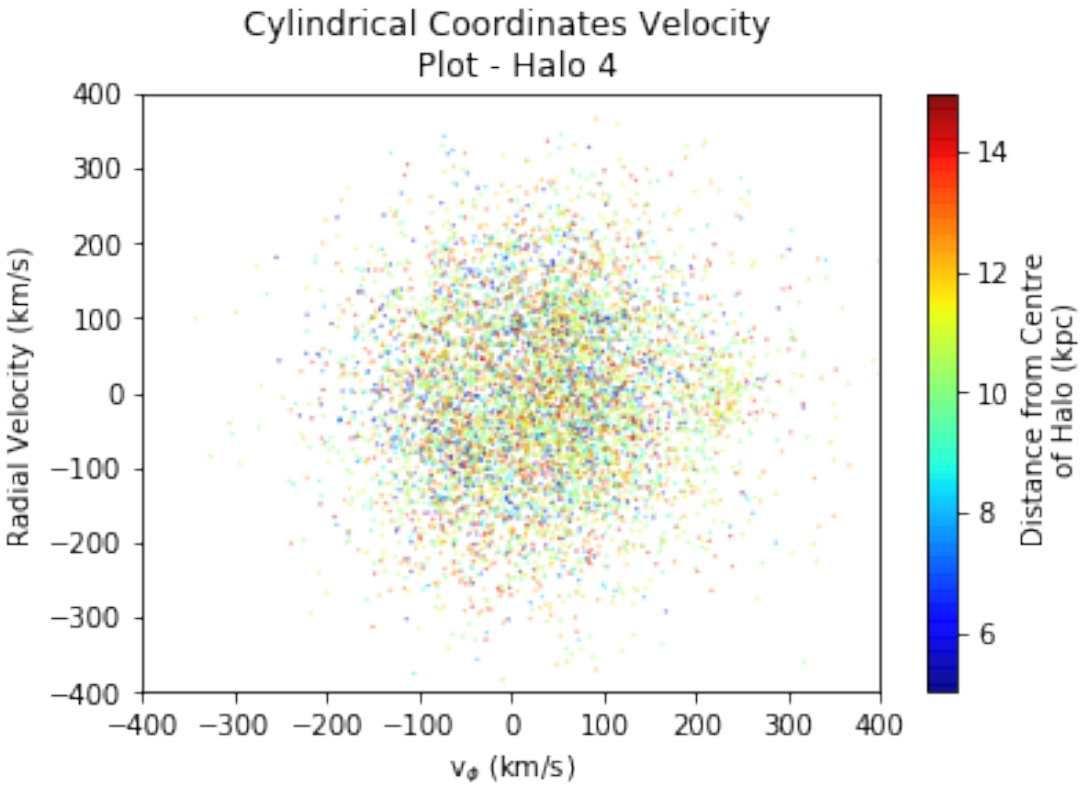

FIGURE 4.6: Velocity plot for halo 4.

4.2 Comparisons Within Haloes

As previously mentioned in Section 2.8, all the statistical results are subject to a 2σ limit of $\log_{10}$ p-value $= -1.3$. This limit was chosen as the results were found to be insensitive to the choice between 2σ and 3σ. Any results below -1.3 are discounted, as they suggest that there is very little similarity between the groups being compared. As the aim of this work is to establish if any groups have similar evolutionary histories to solar type particles, only values above this cut off require any investigation.

This section also refers to results as being 'limited'. This occurs when haloes have undergone a merger, or when one or more of the groups tested contains less than 100 particles. This value was chosen as the distributions of groups with very few particles can be interpreted by the statistical tests as being a better match to a comparison distribution than they are. Any results that include a group with less than 100 particles are given in italics, and any results that involve a halo that has undergone a merger are given in bold.

4.2.1 Statistical Analysis of the Halo-centric Radius

The halo-centric radius distribution of each group within each halo was compared to the distribution of the solar group within the halo. The comparisons were conducted with both the Kolmogorov-Smirnov and Anderson-Darling tests. A result close to 0.000 for the KS test, and a result close to -0.602 for the AD test would suggest a similar distribution of particles between the solar and comparison group, which may suggest that both groups inhabited a region in their host galaxy that was preferential to the development of life. The $\log_{10}$ p-value of each result was taken, and if it was below -1.30, it was discounted. The results above the cut off are outlined in Table 4.4.

TABLE 4.4: Statistically significant results of Kolmogorov-Smirnov and Anderson-Darling tests comparing the distributions of the halo-centric radius within haloes to that of the solar group

Halo	Group	$\log_{10}$ KS p-value	$\log_{10}$ AD p-value
4	*Solar Age Metal Rich*	*-0.325*	*-0.602*
11	*Old Solar Metallicity*	*-0.172*	*-0.602*
11	*Young Metal Rich*	*-0.149*	*-0.602*
11	*Young Solar Metallicity*	*-0.949*	*-1.00*
15	Old Solar Metallicity	-0.381	-0.602
15	*Solar Age Metal Rich*	*-0.404*	*-0.602*
15	Solar Age Metal Poor	-0.692	-0.602
20	Young Solar Metallicity	-1.084	

These tests produced 256 values when combined, but only 15 were found to be above the cut off. This suggests that most of the age-metallicity groups show little to no similarity to the solar group found in their haloes.

The young solar age metallicity group from halo 20 has a Kolmogorov- Smirnov p-value that shows the least similarity to the solar group. This group also does not have a consistent Anderson-Darling result, suggesting that the similarity lies in the body of the distribution, and not in the tails. As Kolmogorov-Smirnov and Anderson-Darling tests both compare the entire distribution, both tests should produce a result above the cut off if the distributions are similar. As this is not the case here, this result is discounted.

This leaves 7 pairs of results above the cut off, with six of these pairs involving a group with either a solar age or metallicity. This suggests that these groups are more similar to the solar group than those without solar properties. However, as this is not consistently found across all haloes, this trend cannot be considered universal.

Whilst these results do provide evidence for similarities between the solar groups and other groups in some haloes, the small sample sizes limit the usefulness of these results. As only four haloes are represented in these results, there are no universal similarities

between the solar groups and any other group. Overall, this suggests that the radius distributions of the solar groups are likely to be drawn from a different distribution to the other groups.

4.2.2 Statistical Analysis of the Mass-Weighted Time

The mass-weighted time provides an indication of how long mass spends in a cluster environment, whilst also accounting for the difference in the masses of individual particles. This means that particles with large amounts of mass in cluster environments can be more easily compared to particles with less mass in cluster environments. The mass-weighted time was analysed using both Kolmogorov-Smirnov and Anderson-Darling tests. These compared the solar distribution to the distributions of the other 8 age metallicity groups and implemented the 2σ cut off. Any values above this cut off are outlined in Table 4.5.

TABLE 4.5: Statistically significant results from the Kolmogorov-Smirnov and Anderson-Darling tests for the mass-weighted time within haloes.

Halo	Group	$\log_{10}$ KS p-Value	$\log_{10}$ AD p-Value
1	*Young Metal Poor*	*-0.003*	*-0.602*
4	*Solar Age Metal Rich*	*-0.225*	*-0.602*
4	*Old Metal Rich*		*-1.209*
6	*Solar Age Metal Poor*	*-1.014*	*-0.891*
11	*Solar Age Metal Poor*		*-1.093*
11	*Young Metal Rich*	*-1.053*	*-0.775*
11	*Young Solar Metallicity*	*-0.181*	*-0.704*
11	*Young Metal Poor*	*-1.013*	*-1.225*
14	*Old Metal Rich*	*-0.772*	
15	Old Solar Metallicity	-0.252	-0.602
15	*Solar Age Metal Rich*	*-0.487*	*-0.602*
24	*Solar Age Metal Rich*	*-0.308*	*-0.602*

As previously mentioned, distributions that show significant similarities are expected to have a pair of results above the threshold, so the un-paired results from Table 4.5 are not considered when determining if there are any trends. This leaves 9 pairs of results.

Eight of these pairs are limited by small sample sizes, reducing the validity of these results. Six pairs refer to groups with either a solar metallicity or solar age, including the only pair not limited by a small sample size. This suggests that populations with a solar property are more similar to the solar population, but as the results only include six pairs from five haloes, this similarity is not consistently found within or across all haloes, meaning this result is unlikely to be indicative of a universal trend.

Like the tests conducted on the halo-centric radius, these results suggest that solar groups within haloes have different distributions of MWT to the other groups. However, groups with solar properties are likely to be more similar than groups without solar properties.

4.2.3 Statistical Analysis of the Number of Supernovae

Potential similarities in the expected number of supernovae between the groups were examined using both Kolmogorov-Smirnov and Anderson- Darling tests. The results indicating significant similarity are outlined in Table 4.6.

TABLE 4.6: Statistically significant results from the Kolmogorov-Smirnov and Anderson-Darling tests conducted on the expected number of supernovae in the first 100 Myrs.

Halo	Group	KS $\log_{10}$ p-Value	AD Test $\log_{10}$ p-Value
1	*Young Metal Poor*	*-0.232*	*-0.602*
4	*Old Metal Rich*	*-0.475*	*-0.602*
4	*Solar Age Metal Rich*	*-0.389*	*-0.602*
6	Old Metal Poor	-0.704	-0.942
6	*Solar Age Metal Poor*	*-1.252*	*-1.231*
6	*Young Metal Rich*	*-1.110*	
6	Young Solar Metallicity	-0.658	-0.635
11	*Old Solar Metallicity*	*-1.070*	
11	*Solar Age Metal Poor*	*-1.061*	*-0.747*
11	*Young Metal Rich*	*-0.172*	*-0.602*
11	*Young Solar Metallicity*	*-0.351*	*-0.602*
14	*Old Metal Rich*	*-0.189*	*-0.602*
15	Old Solar Metallicity	-0.761	-1.161
15	*Solar Age Metal Rich*	*-0.341*	*-0.602*
24	*Solar Age Metal Rich*	*-0.941*	*-0.602*

As only pairs of results are considered when establishing trends, two unpaired results were removed, leaving 13 pairs of results. Eight of these pairs include groups with either a solar age or metallicity, which may suggest that distributions with a solar properties are more similar to the solar group than those without.

This implies that, in these particular haloes, solar type particles and particles with solar properties experience a similar number of supernovae in their early evolutionary histories. As supernovae can both disrupt planet formation and provide the materials necessary to form planets, it is possible that there is an 'ideal' amount of supernovae activity for a star to undergo in its early life. If this is the case, the similarity in the supernovae activity in the early lives of solar particles and particles with solar properties may suggest that both groups are in environments that support planet formation. As the Sun is known to host planets, this similarity may suggest that particles with solar properties are more

likely to host planets than particles without solar properties. However, as mentioned in Section 2.3, these simulations have a resolution limit coarser than that required to model planets, so this possibility cannot be confirmed using these simulations. This could be further examined by future studies using both simulations and observational data. Currently, observational data from the NASA exoplanet archive could be utilised to examine the ages and metallicities of stars known to host exoplanets.

4.2.4 Statistical Analysis of the Time Spent in a Cluster (0-100 Myrs)

Kolmogorov-Smirnov and Anderson-Darling tests were again used to analyse similarities in the amount of time spent in a cluster environment in the first 100 Myrs after particle formation. After employing the previously defined cut off and removing unpaired results, 7 pairs of values remained. The results are outlined in Table 4.7.

TABLE 4.7: Statistically significant results from the Kolmogorov-Smirnov and Anderson-Darling tests carried out on the time spent in a cluster environment during the first 100 Myrs since particle formation.

Halo	Group	KS $\log_{10}$ p-Value	AD $\log_{10}$ p-Value
1	*Young Metal Poor*	*-0.702*	*-0.836*
2	*Solar Age Metal Poor*	*-1.097*	
4	*Old Metal Rich*	*-0.258*	*-0.602*
4	*Solar Age Metal Rich*	*-0.849*	*-1.145*
6	*Young Metal Rich*	*-0.565*	*-0.633*
11	*Solar Age Metal Poor*	*-0.972*	*-0.993*
14	*Old Metal Rich*	*-0.439*	
15	*Solar Age Metal Rich*	*-0.191*	*-0.602*
24	*Old Metal Rich*	*-0.234*	*-0.656*

As with previous statistical tests, very few values were found to be above the cut off, and all of the values above the cut off are limited due to small sample sizes. Both the solar age metal poor group from halo 2 and the old metal rich group from halo 14 lacked an Anderson-Darling result above the cut off. As mentioned in Section 2.8, these tests both analyse the full distribution, but emphasise different regions. Therefore, if the similarities between the distributions are significant, both tests should give a result above the cut off. Alongside this, both of these groups are limited by a small sample size, so neither are considered when determining if there are any trends in the results.

These results do not suggest any significant similarities between the solar distributions and the other distributions in their respective haloes.

4.2.5 Conclusions

Overall, the lack of similarities between the statistical results obtained for the solar type particles and the other groups suggests that solar particles are unique when the results from individual haloes are considered. This lack of similarity is found in the distributions for the halo-centric radius, mass-weighted time, expected number of supernovae, and the time spent in a cluster are examined. This suggests that stars like the Sun have different evolutionary histories to stars with different ages and metallicities. These results support the hypothesis that solar type stars have differences in their evolutionary histories that may make them more likely to form and retain habitable planets.

The results of both the Kolmogorov-Smirnov and Anderson- Darling tests highlight a greater level of similarity between the solar distributions and distributions from groups with either a solar age or metallicity, but this is not consistent across all haloes, and many of these results are limited in some way. It is therefore not possible to conclude if these groups are more similar to solar type particles than other groups.

4.3 Statistical Comparison Between Haloes

As it has now been established that solar particles are likely to have different evolutionary histories to other particles in their haloes, it is important to determine if solar particles from different haloes share similar evolutionary histories. If the solar groups are found to have similar histories, it would suggest that there is a underlying distribution that all solar particles are drawn from, irrespective of the halo they are found in. This would imply that the evolutionary histories of the particles are not due to the initial conditions used for each halo, and are instead an intrinsic property of the particle. To establish if this is the case, Kolmogorov-Smirnov and Anderson-Darling tests were carried out between solar groups from different haloes on the halo-centric radius, mass-weighted time, the expected number of supernovae in the first 100 Myrs, and the amount of time spent in a cluster for the first 100 Myrs.

4.3.1 Comparison of the Halo-centric Radii Between Haloes

To determine if the halo-centric radius distributions for solar groups were similar between haloes, both Kolmogorov-Smirnov and Anderson-Darling tests were carried out. To account for the different radii of each halo, the values from each halo were divided by the maximum radius of the halo, giving a radius as a fraction of the maximum halo radius. Each solar distribution was compared to all the other solar distributions, giving

120 values. We applied the same 2σ cut off of -1.3 as before as the results are insensitive to this. The results are detailed in Table 4.8.

TABLE 4.8: Statistically significant results from the Kolmogorov-Smirnov and Anderson-Darling tests conducted on the halo-centric radius between solar groups from different haloes.

Halo Numbers	KS $\log_{10}$ p-Value	AD $\log_{10}$ p-Value
11 and 4	*-1.291*	
11 and 12	*-0.362*	*-0.633*
20 and 4	-0.709	

Only one of the Kolmogorov-Smirnov test had a consistent Anderson- Darling test value. As distributions that show significant similarities are expected to have results above the 2σ cut off for both tests, any results with only one value are discarded. This leaves only one result, which includes halo 11. As this halo contains less than 100 solar particles, the statistical results have the potential to be skewed. Therefore, it is not possible to accurately identify any trends using this result. Overall, the lack of results above the 2σ cut off suggests that there is little similarity between solar groups from different haloes.

4.3.2 Comparison of the Mass-Weighted Time Between Haloes

To compare the mass-weighted time of solar groups from different haloes, both Kolmogorov-Smirnov and Anderson-Darling tests were used. The Kolmogorov-Smirnov test yielded 8 $\log_{10}$ p-values that were above the cut off, and the Anderson-Darling test gave 6 values. These are outlined in Table 4.9.

TABLE 4.9: Statistically significant results from the Kolmogorov-Smirnov test conducted on the mass weighted time between solar groups from different haloes.

First Halo Number	Second Halo Number	Kolmogorov-Smirnov log_{10} p-value	Anderson-Darling log_{10} p-value
11	*2*	*-1.127*	
11	3	*-0.665*	
11	4	*-0.929*	*-0.833*
11	10	*-0.458*	*-0.826*
11	12	*-0.664*	*-1.151*
11	20	*-0.181*	*-0.602*
20	3	-0.924	-1.175
22	15	**-1.020**	**-1.262**

The comparisons between haloes 11 and 2, and 11 and 3 lack an Anderson-Darling value, suggesting any similarities are limited to the main body of the distributions. Both pairs also have limited sample sizes, which has the potential to impact the results of the tests. Therefore, these results are not considered when identifying any trends.

This leaves 6 pairs of results. However, 4 of these are limited by small sample sizes, and one is impacted by a major merger, limiting the usefulness of these results. This leaves 1 result that is not limited by either a merger or a small group. As there were 120 possible combinations, the small number of results suggests that there is little similarity between the mass-weighted times of solar particles from different haloes. This suggests that the distribution of mass-weighted time of solar particles is impacted by the conditions they experience, which are determined by the initial conditions of the halo they reside in.

4.3.3 Comparison of the Expected Number of Supernovae Between Haloes

Kolmogorov-Smirnov and Anderson-Darling tests were conducted on the expected number of supernovae between the solar groups in each halo. The results are outlined in Table 4.10.

TABLE 4.10: Statistically significant results from the Kolmogorov-Smirnov and Anderson-Darling tests conducted on the number of supernovae experienced in the first 100Myrs between solar groups from different haloes.

Halo Number Number	Halo Number Number	Kolmogorov-Smirnov $\log_{10}$ p-Value	Anderson-Darling $\log_{10}$ p-Value
1	*11*	*-1.222*	*-0.872*
11	20	*-0.661*	*-1.040*
11	24	*-0.395*	*-0.602*
15	**22**	**-1.020**	

The comparison between haloes 15 and 22 lacks an Anderson-Darling value, and the Kolmogorov-Smirnov value is limited due to the major merger in halo 22. Therefore, this result is not considered when establishing trends.

This leave 3 pairs of values which are all limited by small sample sizes. These results suggest that solar type particles in different haloes experience different numbers of supernovae in their early lives.

As solar type particles are representative of the Sun, and the Sun is known to host a planetary system, this result could be taken to suggest that supernovae events do not have a significant impact on the evolution of stars and the formation of planets. However, this conclusion would rely on the assumption that all solar particles in all haloes represented stars that hosted planets. The resolution of these simulations does not allow for planetary systems to be resolved, so it is not possible to determine if all solar particles have planetary systems, and how much impact supernovae have had on these systems if they were to exist. The number of planetary systems hosted by solar

type stars and the impact of supernovae on these systems could be determined in future work using simulations and observational data.

These results suggest that solar type particles experience varying distributions of supernovae in their early lives, meaning some particles may find themselves in environments that are significantly more hostile to planetary formation and evolution than others. As there is no clear reason for certain particles to be in more hostile environments, it is likely that the variation is due to the initial conditions used in individual haloes.

It is important to note that these simulations do not resolve individual stars and supernovae, so variables such as the distance between a star and a supernova are not accounted for in these results. It would therefore be useful for future work to factor in these variables.

4.3.4 Comparison of the Amount of Time Spent in a Cluster Environment Between Haloes

When Anderson-Darling tests were conducted on the time spent in clusters for different solar groups, there were no values above the cut off. The Kolmogorov-Smirnov tests produced three values above the -1.3 cut off, and these are detailed in Table 4.11.

TABLE 4.11: Statistically significant results from the Kolmogorov-Smirnov tests conducted on the time spent in a cluster environment between solar groups from different haloes.

Halo Number	Halo Number	Kolmogorov-Smirnov $\log_{10}$ p-Value
11	15	*-0.947*
11	20	*-0.844*
15	**22**	**-1.020**

As none of these results is part of a pair, all of these can be discarded, meaning there is no available evidence to suggest any similarities between the solar groups from different haloes when considering the amount of time spend in a cluster environment. In addition to this, two of these results include halo 11, which only contains 57 solar particles, and the other contains halo 22, which has undergone a merger, meaning these results are of limited use.

4.3.5 Conclusions

The statistical tests conducted between solar groups from different haloes do not show any strong similarities between the solar groups. This suggests that solar type particles

from different haloes have different evolutionary histories, which are likely to be the result of different initial conditions.

4.4 Establishing Links Between Metallicity and Evolutionary History

It has now been established that solar type particles have different evolutionary histories to particles from other groups, but do not share a common evolutionary history. However, as the simulations cover the evolution of particles from the Big Bang to the present day, it is also possible to determine if there are any links between particle metallicity and their evolutionary histories. The following sections examine trends found between the metallicity of star particles and the initial fraction of mass found in a cluster environment, the amount of time spent in a cluster environment, the disruption time of cluster environments, and the number of supernovae experienced.

4.4.1 Link Between the Initial Percentage of Mass in a Cluster Environment and Metallicity

To determine if there is a link between the initial percentage of mass in a cluster environment and the metallicity of the particles, the mean initial percentage of mass in a cluster environment for each group from each halo was analysed. Table 4.12 details the results, with the standard error on the mean included.

When tracking the change in the fraction of mass in a cluster environment for individual particles, it became apparent that particles classified as metal rich had more of their initial mass in a cluster environment on average. Particles classified as metal poor often had the lowest mean initial fraction of mass in a cluster environment, and solar metallicity particles sat between metal rich and metal poor particles. In part, this is due to higher metallicity environments having higher cooling efficiencies. This would usually cause lower temperatures in these environments, but the EMP simulations have a temperature floor of 10^4K. This means that fast cooling, high metallicity particles will eventually reach a state of constant temperature. As pressure can be approximated as the product of density and temperature, any changes in pressure must be due to changes in density. An increase in pressure will therefore be caused by an increase in density, and these denser regions will have a higher star formation efficiency due to shorter free fall times, leading to a higher initial fraction of mass in a cluster environment.

TABLE 4.12: The initial percentage of mass in a cluster environment for all groups in all haloes.

Halo	Percentage of Initial Mass in Cluster Environment							
Number	Old Rich	Solar Rich	Young Rich	Old Solar	Solar	Young Solar	Old Poor	Solar Poor	Young Poor
1	37.5 ±1.2	35.6 ±0.8	40.2 ±0.1	19.5 ±0.1	17.6 ±0.2	14.5 ±0.1	15.6 ±0.0	n/a	*11.2* ±2.8
2	47.5 ±0.1	50.4 ±0.2	64.2 ±0.0	31.4 ±0.1	*19.5* ±0.9	30.5 ±0.8	16.7 ±0.0	*8.61* ±1.8	8.63 ±0.1
3	26.1 ±0.6	33.6 ±0.3	49.7 ±0.0	23.6 ±0.1	28.0 ±0.1	29.4 ±0.1	20.2 ±0.1	n/a	n/a
4	*26.4* ±1.5	*27.8* ±3.1	37.3 ±0.6	22.1 ±0.1	23.3 ±0.3	20.6 ±0.1	15.1 ±0.1	*7.93* ±1.8	8.29 ±0.3
5	52.8 ±0.1	56.3 ±0.1	58.4 ±0.0	40.2 ±0.1	40.4 ±0.3	37.6 ±0.2	23.1 ±0.2	n/a	n/a
6	*22.3* ±1.8	*26.7* ±4.7	*20.1* ±2.6	17.9 ±0.1	13.0 ±0.5	12.6 ±0.1	15.0 ±0.1	*3.5*±0.8	7.9±0.3
10	51.2 ±0.1	56.1 ±0.1	61.5 ±0.0	33.5 ±0.1	27.4 ±0.2	15.5 ±0.1	15.6 ±0.0	10.9 ±0.4	9.6 ±0.1
11	n/a	n/a	*19.4* ±2.5	17.1 ±0.2	19.5 ±1.8	18.3 ±0.1	12.8 ±0.0	17.2 ±0.2	14.6 ±0.1
12	49.3 ±0.0	56.7 ±0.0	58.7 ±0.0	24.8 ±0.1	28.8 ±0.3	15.7 ±0.0	14.5 ±0.0	11.1 ±0.2	9.85 ±0.1
13	46.6 ±0.0	57.3 ±0.1	61.1 ±0.0	21.1 ±0.0	14.4 ±0.2	17.1 ±0.1	16.0 ±0.0	8.4 ±0.3	10.1 ±0.1
14	*15.9* ±1.6	n/a	39.9 ±0.1	17.3 ±0.1	15.8 ±0.2	18.3 ±0.0	14.1 ±0.0	11.3 ±0.4	12.5 ±0.1
15	n/a	*7.4* ±0.0	30.3 ±0.8	13.3 ±0.3	13.9 ±1.0	21.9 ±0.1	15.2 ±0.1	19.6 ±0.2	18.3 ±0.1
20	n/a	n/a	n/a	17.4 ±0.2	20.0 ±0.7	17.3 ±0.2	13.3 ±0.1	13.2 ±0.3	10.8 ±0.1
22	**41.7** ±0.3	**43.9** ±0.1	**49.4** ±0.0	**18.7** ±0.1	**15.2** ±0.1	**15.4** ±0.0	**10.9** ±0.0	**6.04** ±0.2	**0.603** ±0.0
23	54.0 ±0.0	59.0 ±0.1	61.6 ±0.0	34.9 ±0.1	49.8 ±0.4	46.3 ±0.2	23.6 ±0.1	n/a	n/a
24	*24.9* ±2.1	*26.0*±4.7	40.9 ±0.1	19.8 ±0.1	19.8 ±0.1	18.4 ±0.1	13.8 ±0.0	10.1 ±0.5	10.2 ±0.2

From Table 4.12, it is clear that haloes 2, 4, 5, 12, 22, 23, and 24 all broadly follow the proposed trend. However, halo 2 only contains 76 solar particles, the old and solar age metal rich and solar age metal poor groups in halo 4 contain less than 100 particles, haloes 5 and 23 have no solar age metal poor or young metal poor particles, halo 22 contains particles from a merger, and halo 24 has fewer than 100 old metal rich and solar age metal rich particles. This means that the only halo that fully follows the trend with no merger particles, missing groups, or statistically limited results is halo 12. This, and the evolution of the fraction of mass in a cluster environment is further illustrated by Figures 4.7 and 4.8. Whilst Figure 4.7 suggests an exponential law governing the mass loss, Figure 4.8 shows that this is not the case at later times.

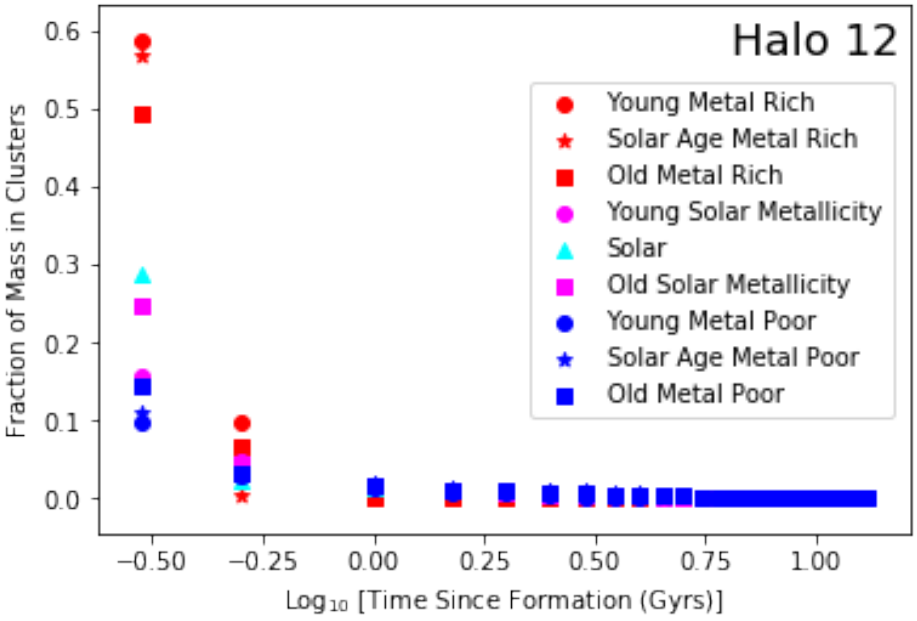

FIGURE 4.7: Scatter plot showing the mean fraction of mass in a cluster environment for each age-metallicity group in Halo 12 as a function of time.

Haloes 1, 6, 10, and 13 have metal rich groups containing the largest fraction of mass in a cluster environment, but do not follow the proposed trend with solar metallicity and metal poor groups. The metal rich groups in halo 6 all contain fewer than 100 particles, which limits the statistical robustness of this result.

Halo 20 contains no metal rich particles. However, the solar metallicity groups do contain a higher fraction of mass in a cluster environment compared to the metal poor groups, showing that the particles that are present do still follow the suggested trend.

Halo 14 follows the proposed trend with the exception of old metal rich particles. However, this groups consists of only 16 particles, limiting the statistical reliability of the group. If old metal rich particles are removed from this halo, the proposed trend is followed.

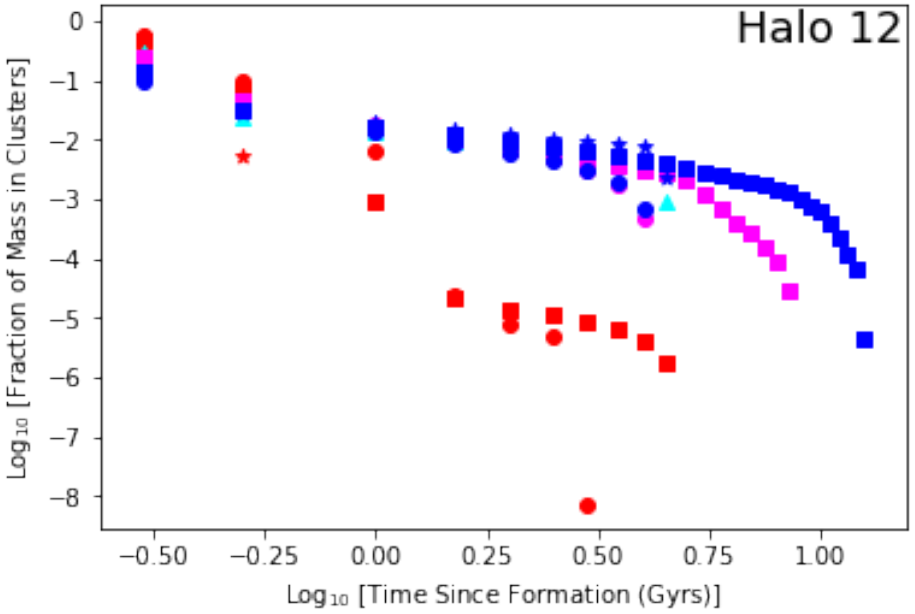

FIGURE 4.8: Scatter plot showing the mean fraction of mass in a cluster environment for each age-metallicity group in Halo 12 on a logged axis.

Table 4.12 also highlights that haloes 3, 11 and 15 do not follow the proposed trend.

Out of the 16 haloes analysed, 7 follow the proposed trend, 6 partially follow the trend, and 3 do not, providing evidence for a link between the metallicity of a particle and the initial fraction of mass in a cluster environment.

Taking star particles as an analogue for stars, these results show that stars with higher metallicities spend more time in a cluster environment. This exposes them to higher stellar densities, and increases the risk that they will be impacted by supernovae events. If these stars have protoplanetary disks or young planets, spending a prolonged period of time in a cluster environment will increase the risk that they will be negatively impacted by the hostile conditions found in their early environments.

4.4.2 Trends in Cluster Disruption Time as a Function of Metallicity for Old Particles

When discussing the timescales for cluster disruption, it is important to note that the age cuts used in this work will impact the results. As young and solar age particles have an imposed upper age limit, these groups may not be disrupted by the end of the simulation. However, for old particles, the only limit is the maximum age of the simulation. This makes trends including young and solar age particles hard to determine, but trends in older particles are easier to find. We therefore separate the particles into age groups and discuss them separately.

To more accurately pinpoint the amount of time required for clusters to disrupt, the fraction of mass in a cluster environment was interpolated with time steps of 0.5 Gyrs. The average fraction of mass in a cluster environment was then taken at each time step for each of the 3 groups being examined. This was done independently for each halo. The cluster environment in a particle was considered to be 'disrupted' when there was no mass left in a cluster environment. Table 4.13 lists the time step at which the amount of mass in a cluster environment first reached 0. As the interpolation was carried out at intervals of 0.5 Gyrs, the particle could have been fully disrupted in the last 0.5 Gyrs, meaning each value has an error of -0.5 Gyrs

TABLE 4.13: Details for the cluster disruption times for old star particles. In cases where there is not complete cluster disruption, the value used is the percentage of the initial mass remaining at the end of the simulation. Values in italics are samples containing less than 100 particles, and values in bold signify a merger.

Halo Number	Disruption Time (Gyrs) or Initial Mass Remaining (%)		
	Old Rich	Old Solar	Old Poor
1	5.0	9.5	13.0
2	3.5	9.0	13.5
3	2.5	7.0	11.0
4	*1.5*	10.5	13.5
5	6.0	7.5	9.5
6	*4.5*	0.094%	13.5
10	5.5	10.0	13.0
11	n/a	9.0	13.5
12	5.0	9.0	13.0
13	7.5	9.5	13.5
14	*1.5*	10.5	13.0
15	n/a	9.0	13.0
20	n/a	8.0	13.5
22	**7.0**	**10.0**	**13.5**
23	4.5	10.0	11.5
24	*1.5*	8.0	13.5

In general, we find that old metal poor particles take longer to disrupt than old solar metallicity particles, which in turn take longer to disrupt than their old metal rich counterparts. This trend is followed by 15 out of the 16 haloes. It is important to note that haloes 11, 15, and 20 do not have any old metal rich particles, haloes 4, 14, and 24 have less than 100 old metal rich particles, and halo 22 has undergone a merger, limiting the statistical robustness of the results from these haloes. The only halo that doesn't follow this trend is halo 6, which does not see old solar metallicity particles undergo cluster disruption.

The link between metallicity and disruption time is partly due to feedback mechanisms from stellar evolution. Metal rich stars have strong stellar winds (Amard & Matt, 2020; Meynet et al., 2006) and are more likely to end their lives as a supernova compared to

stars with lower metallicities (Langer, 2012). Therefore, metal rich clusters experience more aggressive feedback from both winds and supernovae compared to metal poor clusters (Ventura et al., 2020). This feedback will expel mass from the cluster, shortening the cluster lifetime (Krause, Martin G. H. et al., 2016). In addition to this, intense feedback can prevent the formation of new stars by disrupting the remaining gas, so the stellar mass that is lost is not being replaced, and the cluster eventually loses all its mass.

4.4.3 Trends in Cluster Disruption Time as a Function of Metallicity for Solar Age Particles

Solar age particles have an upper age limit of 4.7 Gyrs and the data has been interpolated in 0.5 Gyr intervals. This means that any particles with mass still present in a cluster after 4.5 Gyrs are considered to have failed to disrupt. However, this does not mean that these particles will not disrupt in the future, but the current EMP simulations do not evolve the particles beyond the present time, so it is not known if or when these particles disrupt.

The method used to calculate the disruption time for solar age particles is the same as that outlined in Section 4.4.2, and the error on all results is -0.5 Gyrs. The results are shown in Table 4.14.

TABLE 4.14: Details for the mean cluster disruption times for solar age star particles in haloes. Values in italics are samples containing less than 100 particles.

Halo	Disruption Time (Gyrs) or Initial Mass Remaining (%)		
Number	Solar Rich	Solar Age and Metallicity	Solar Poor
1	0.455%	2.163%	n/a
2	1.5	*2.5*	*0.5*
3	2.5	0.067%	n/a
4	*1.5*	0.817%	*1.5*
5	0.5	4.5	n/a
6	*1.0*	3.119%	*1.0*
10	1.0	0.617%	2.040%
11	n/a	*7.101%*	0.018%
12	1.0	0.331%	2.044%
13	0.0002%	1.714%	1.881%
14	n/a	0.539%	0.009%
15	*0.5*	0.202%	0.007%
20	n/a	1.796%	2.188%
22	3.5	2.075%	4.0
23	1.0	1.0	n/a
24	*0.5*	0.488%	*4.5*

No haloes have particles in all three groups and follow the trend identified in Section 4.4.2. Haloes 3 and 5 are both lacking solar age metal poor particles, but the other groups in the haloes do follow the proposed trend.

Both haloes 22 and 24 have metal rich particles disrupting before metal poor particles, which follows the proposed trend of disruption time increasing as metallicity decreases. However, solar metallicity particles do not disrupt, meaning this halo does not follow the trend.

Solar age metal poor particles disrupt first in halo 2, followed by solar age metal rich and solar particles. However, both the solar and solar age metal poor groups have less than 100 particles, so it is possible that this result has been skewed by small sample sizes.

Both haloes 4 and 6 have the equal disruption times for solar age metal poor and solar age metal rich particles. Solar metallicity particles do not disrupt in either halo, meaning they do not follow the proposed trend. However, the only group in each halo that is not limited by a small sample size are solar particles, meaning these results may have been skewed by small sample sizes.

Haloes 1 and 23 lacks any solar age metal poor particles, and both solar age metal rich and solar particles either take the same amount of time to disrupt, or do not disrupt. Due to the errors on the measurements, it is possible that halo 23 does follow the trend. It is not currently possible to determine if halo 1 follows the proposed trend, as the EMP simulations do not evolve past redshift 0.

No groups in halo 13 disrupt, meaning it is not possible to determine if this halo may follow the trend.

Haloes 10, 11, 12, 14, 15, and 20 all have solar and solar age metal poor groups which do not undergo disruption. This again means that it is not possible to determine if these haloes will follow the trend in the future, but it cannot be ruled out.

Due to the upper age limit imposed on the particles, it was expected that more groups would fail to disrupt. However, if the solar age particles were following the trend identified in Section 4.4.2, it is expected that any solar groups that fail to disrupt would be paired with a metal poor group that either contains no particles, or also fails to disrupt. Young metal rich groups that fail to disrupt are expected to be part of a trio of groups that either fail to disrupt or lack particles. In both of these scenarios, it would be possible to conclude that a given halo may eventually follow the proposed trend if evolution was allowed to continue. However, 4, 6, 22, and 24 have solar particles failing

to disrupt, but solar poor particles disrupting. It should be noted that the solar poor groups in haloes 4, 6, and 24 are limited by small sample sizes.

The results in Table 4.14, show a surprisingly high number of solar groups failing to disrupt. If these results were following the proposed trend, it is expected that metal poor groups would make up the majority of the groups that failed to disrupt, followed by solar metallicity groups. However, these results show that 13 out of the 16 solar groups failed to disrupt, whilst the same is only true for 7 metal poor groups. This suggests that particles with both a solar age and metallicity retain more of their mass in a cluster environment that particles of a similar age but different metallicities.

To determine the disruption times of groups that are yet to disrupt, the EMP simulations would need to be evolved in to the future.

4.4.4 Trends in Cluster Disruption Time as a Function of Metallicity for Young Particles

As with the solar age particles, young particles are subjected to an upper age limit. In this case, the upper age limit is 4.3 Gyrs. Coupled with the interpolation used, this means that any particles that still have cluster mass after 4.0 Gyrs are classified as not disrupted. The method used is that defined in Section 4.4.2, and all results have an error of -0.5 Gyrs. The results are shown in Table 4.15.

TABLE 4.15: Details for the mean cluster disruption times for young star particles in haloes. Values in italics are samples containing less than 100 particles.

Halo	Disruption Time (Gyrs) or Initial Mass Remaining (%)		
Number	Young Rich	Young Solar Metallicity	Young Poor
1	0.009%	0.715%	*0.5*
2	1.5	1.5	4.0
3	0.001%	0.023%	n/a
4	4.0	0.169%	*3.5*
5	3.0	3.5	n/a
6	*0.019%*	0.466%	*3.5*
10	2.0	0.159%	0.717%
11	*0.308%*	0.018%	2.666%
12	3.5	0.310%	0.728%
13	4.0	0.187%	0.183%
14	3.0	0.059%	0.132%
15	3.5	0.007%	3.456%
20	n/a	0.002%	0.359%
22	0.001%	0.296%	0.160%
23	1.5	2.0	n/a
24	4.0	0.132%	4.0

As with solar age particles, the expected trend was that disruption time would increase as metallicity decreased. No haloes have particles in all three groups and also follow the proposed trend. Haloes 5 and 23 both lack young metal poor particles, but their young metal rich populations disrupt before their young solar metallicity populations, meaning they follow the proposed trend where possible.

Halo 2 has young metal rich and young solar metallicity both disrupting after 1.5 Gyrs. However, when the errors are taken into account, it is possible that young solar metallicity particles are disrupting after young metal rich particles. As young metal poor particles in this halo disrupt after 4 Gyrs, it is possible that this halo does follow the predicted trend.

Haloes 1 and 6 both have young metal poor particles disrupting, but neither halo sees any disruption in young metal rich or young solar metallicity particles. However, the young metal poor groups in both haloes are limited by small sample sizes, which may have caused a reduction in the disruption time. Neither of these haloes can be considered to follow the proposed trend.

Similarly, halo 4 has no disruption in young solar metallicity particles, but has young metal rich particles disrupting after 4 Gyrs and young metal poor particles disrupting after 3.5 Gyrs. Again, the young metal poor group is limited by a small sample size, but due to the lack of disruption in young solar metallicity particles, this halo cannot be considered to follow the proposed trend.

Young solar metallicity particles also do not disrupt in halo 24. The young metal rich and young metal poor particles in this halo both disrupt at 4 Gyrs, meaning this halo also does not follow the proposed trend.

Haloes 10, 12, 13, 14, and 15 all have young metal rich groups that undergo disruption, but no disruption in young solar metallicity or young metal poor groups. This means that these haloes may follow the proposed trend if they are allowed to evolve further, but as this has not been done, it cannot be confirmed if this will be the case.

Haloes 3, 11, 20, and 22 do not have any young groups undergoing disruption, so it is not possible to determine if these haloes will follow the proposed trend.

Ideally, a future study would be conducted which allowed these simulations to be evolved into the future. Whilst this would prevent the simulations from being calibrated or compared to observations, it may provide some insight into whether these galaxies do eventually follow the disruption trend seen in old clusters.

4.4.5 Trends In Metallicity and The Amount of Time Spent in a Cluster Environment

Whilst cluster disruption provides insight into when all the mass has left a cluster environment, it does not determine how long an average star may remain in a cluster environment. In order to determine this, the expectation value for the amount of time spent in a cluster environment in the first 100 Myrs after particle formation was calculated. This was done by multiplying the interpolated fraction of mass in a cluster environment by the corresponding interpolated time interval. The sum of these was then taken for each particle. The values from corresponding groups from all 16 haloes were then combined, and the mean, median, maximum, minimum, and standard deviation for each group was taken. These values are presented in Table 4.16, with the mean value also including the standard error on the mean. A young metal poor particle was identified as having spent a very short amount of time in a cluster environment, which skewed the results for the minimum, mean, median, and standard deviation. Therefore, these values were calculated twice for this group: once with the outlier included, and once without. The values in the brackets are those obtained when this outlier was removed.

TABLE 4.16: Statistical analysis of the expectation value for the amount of time spent in a cluster environment (0-100 Myrs). Values calculated with the outlier removed are in brackets.

Group	Minimum (Myrs)	Maximum (Myrs)	Median (Myrs)	Mean (Myrs)	Standard Deviation
Young Metal Rich	1.218	8.281	6.526	5.868 ±0.549	2.127
Solar Age Metal Rich	1.160	8.272	6.188	5.906 ±0.605	2.180
Old Metal Rich	1.953	7.592	3.485	4.468 ±0.525	1.892
Young Solar Metallicity	0.222	4.786	1.812	2.223 ±0.329	1.315
Solar Age and Metallicity	1.216	7.793	2.257	2.820 ±0.412	1.648
Old Solar Metallicity	1.208	5.358	2.343	2.7874 ±0.323	1.291
Young Metal Poor	0.076	1.515	1.041	1.020 ±0.098	0.355
	(0.764)	(1.515)	(1.061)	(1.099 ±0.066)	(0.237)
Solar Age Metal Poor	0.349	1.859	1.210	1.205 ±0.121	0.418
Old Metal Poor	1.465	4.105	1.886	2.165 ±0.165	0.658

From Table 4.16, it can be seen that the mean, median, and standard deviation all decrease as metallicity decreases. This suggests that metal rich particles spend, on average, more time in a cluster environment during the first 100 Myrs of their lives. However, there is a larger spread in the values obtained for metal rich particles, suggesting that the amount of time these particles spend in a cluster environment is more varied.

4.4.6 Trends In Metallicity and the Number of Supernova Experienced

The potential number of supernovae experienced by a star particle are outlined in Table 4.17, alongside the standard error on the mean. As with the expected amount of time spent in a cluster, the outlier found in the young metal poor group impacted the results. Therefore, the values in the brackets for this group represent those calculated when the outlier was removed.

TABLE 4.17: Statistical analysis of the expected number of supernovae experienced in the first 100 Myrs since formation for each group across all haloes.

Group	Minimum	Maximum	Median	Mean	Standard Deviation
Young Metal Rich	797	7902	4989	4641 ±575	2228
Solar Age Metal Rich	114	6790	3954	4043 ±606	2186
Old Metal Rich	444	5688	4309	3493 ±513	1849
Young Solar Metallicity	1254	5375	2343	2566 ±241	965
Solar Age and Metallicity	788	5513	2201	2418 ±317	1269
Old Solar Metallicity	1200	4563	2549	2590 ±204	817
Young Metal Poor	35	2557	1241	1213 ±208	751
	(317)	(2557)	(1245)	(1311 ±201)	(697)
Solar Age Metal Poor	30	3736	785	1225 ±329	1141
Old Metal Poor	1585	3765	2083	2271 ±153	611

From Table 4.17, it was again apparent that the median, mean, and standard deviation values all decreased as the metallicity decreased. This suggests that metal rich particles are likely to experience more supernovae in the first 100 Myrs of their lives than solar metallicity or metal poor particles. This is not unexpected, as it has already been established that metal rich particles spend more time in a cluster environment during this time period, which increases the likelihood that these particles will be in the vicinity of a high mass star that will go supernova.

It should be noted that the values in Table 4.17 are the averages for each age-metallicity group. This means that the larger clusters in each group, which experience significantly more supernovae events, are included alongside smaller clusters with few or no supernovae events. This means that this average does not consider the relative sizes of the clusters, so these values are likely to be more representative of larger clusters with many supernovae events. As the mass distribution between clusters has not been included in this work, it was not possible to perform a weighted mean, but this would likely improve the accuracy of the results.

4.4.7 Correlation Between Time in a Cluster Environment and the Expected Number of Supernovae

After establishing a clear link between metallicity and both the time spent in a cluster environment and the number of supernovae, it was decided to plot these together to establish if there was a clear correlation. Figure 4.9 shows a positive correlation between the two variables, with time spent in a cluster environment increasing as the total number of supernovae increase. It is also notable that the spread in the number of supernovae experienced increases with metallicity, with the range being 7787 for metal rich particles, 4725 for solar metallicity particles, and 3735 for metal poor particles. This suggests that the number of supernova experienced by metal poor particles is somewhat more predictable than the number experienced by metal rich, mainly because metal poor particles experience fewer supernovae.

It is important to note that both the time spent in a cluster environment and the number of supernovae experienced are dependent on the fraction of mass in a cluster environment. However, when these variables were independently analysed, they both showed an increase with metallicty, so the correlation seen in the scatter plot does not negate any previous results. Ideally, a partial Spearman rank test would be carried out to confirm that the correlation seen in the lower left panel is correct, the different initial conditions of each of the 16 haloes represented in this plot makes that impossible.

4.4.8 Conclusions

These result highlighted a merger in halo 22, which has provided useful context for the statistical results obtained from this halo.

We also find that lower metallicity particles spend longer in cluster environments than particles of similar ages but higher metallicities. As cluster environments are hostile to planet formation, this result suggests that planets are less likely to be found around these stars, and any planets that are found may have formed in hostile conditions, impacting their habitability.

The initial fraction of mass in a cluster environment is found to generally increase as metallicity increases. This suggests that high metallicity particles are likely to have more stellar mass in a cluster environment, which is not conducive to planet formation or habitability. As star particles are being used as a proxy for stars, this suggests that higher metallicity stars may spend their early lives in hostile environments, meaning that any protoplanetary disks that form may be destroyed.

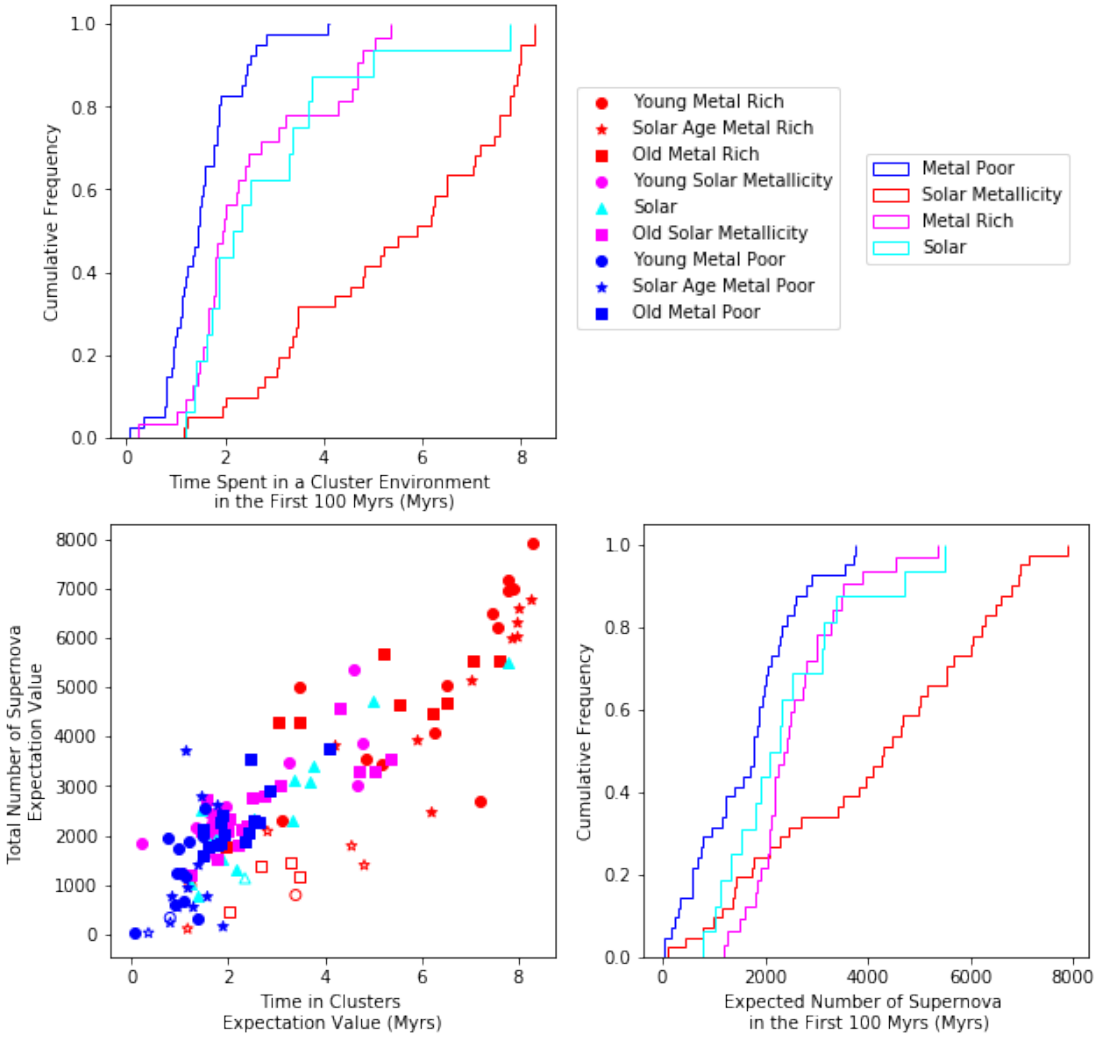

FIGURE 4.9: Upper left panel shows the cumulative frequency plots for the amount of time spent in a cluster environment in the first 100 Myrs. Lower right panel shows the cumulative frequency plots for the expected number of supernovae experienced in the first 100 Myrs. Bottom left panel is a scatter plot showing the correlation between the metallicity and the expected number of supernovae.

When the cluster disruption time is examined in old particles, we find that the time taken for clusters to fully disrupt increases as metallicity decreases. This would suggest that old metal poor stars spend more time in a cluster environment. However, the results from the investigation into the initial fraction of mass in a cluster environment show that old metal poor particles have less mass in a cluster environment, which suggests that fewer stars inhabit old metal poor clusters. Combined, this suggests that old metal poor clusters will contain fewer stars, but that these stars remain in this environment for prolonged periods, potentially limiting their ability to form habitable planets due to the hostile environments they face.

When solar aged particles were examined, significantly more groups failed to disrupt, meaning they retained at least some mass in a cluster environment until the end of

the simulations. Two solar age metal rich groups failed to disrupt, but both of these had less than 0.5% of their initial mass still in a cluster environment at the end of the simulations. Seven solar age metal poor groups failed to disrupt, and all of these has retained less than 3% of their initial cluster mass. Of the 16 solar groups, 13 failed to disrupt. Twelve of these had retained less than 4% of their initial mass, with the exception being halo 11. In halo 11, 7.101% of the initial cluster mass is still in a cluster environment at the end of the simulations. This is around 10-20 times larger than most other groups that fail to disrupt. However, halo 11 contains less than 100 solar particles, so it is possible that a larger cluster has skewed this result. Therefore, when the mean percentage of initial mass left in a cluster environment is considered, the result from the solar group from halo 11 is not included. This leaves solar age metal poor with an average of 1.170% of their initial cluster mass still present in cluster environments at the end of the simulation, compared to 1.161% for solar particles. This suggests that, were the simulations to continue evolving, some galaxies may see solar particles disrupting slightly earlier than their metal poor counterparts, repeating the result seen in old particles.

When young particles were analysed, it was discovered that 5 haloes had young metal rich particles that failed to disrupt, but that each of these haloes had retained less than 0.4% of their initial mass. Thirteen young solar metallicity groups failed to disrupt and 8 young metal poor groups failed to disrupt. However, the average percentage of initial mass remaining in a cluster environment was 0.194% for solar metallicity particles, and 1.050% for metal poor particles. This suggests that metal poor particles have disrupted to a lesser extent than solar metallicity particles, which may suggest that young particles may replicate the trend seen in old particles if they are allowed to continue evolving.

The results from both the solar age and young analysis fail to replicate the results seen in old particles. This can be at least partially explained by the shorter timescales used. The limited time available to these clusters may prevent them from disrupting, which in turn prevents any conclusions being drawn about possible trends.

All particles that failed to disrupt had very little mass remaining in cluster environments, suggesting the processes causing disruption are efficient when combined. It may also suggests that these particles will undergo complete disruption relatively soon, but this cannot be confirmed in this work.

It was also determined that the number of supernovae experienced in the first 100 Myrs since particle formation decreased as metallicity decreased. This implies that metal rich particles will experience a greater number of supernovae in this time period. Based on previous results, it has already been established that these particles spend more time in a cluster environment, which would suggest that they are more likely to be near an O

or B type star that will go supernova at the end of its life, meaning this result was not unexpected.

Finally, a clear positive correlation between the amount of time spent in a cluster environment, the expectation number for the total number of supernovae, and metallicity was observed. This result was not unexpected, as it was previously established that metal rich particles spend more time in a cluster environment and experience more supernovae. However, there had previously been no indication that this result would be roughly linear when considered across all 16 haloes. This suggests that this result is due to one or more fundamental properties of the simulations.

Chapter 5

Results from the Galaxy Analysis

As the Milky Way is the most observationally studied galaxy, the analysis was run again to compare only the Milky-Way-mass galaxy within each halo. The galaxies used in this work have halo masses of between 0.81 - 2.25 $10^{12} M_{\odot}$. They are referred to using the notation M_{200}, which is defined as the mass enclosed within a sphere with an average density 200 times that of the critical density. This is compared to the Milky Way value of 1.08 $^{+0.20}_{-0.14}$ $10^{12} M_{\odot}$ (Reina-Campos et al., 2022a). For the detailed masses of each galaxy, the reader is directed to the values listed under the Constant SFE columns of Table 2 in (Reina-Campos et al., 2022a).

To isolate the particles within each galaxy, the scale height of each galaxy was calculated (see Section 3.1 for details), and only particles within the scale height were considered. This reduced the number of particles within each halo, meaning some age-metallicity groups now had fewer than 100 particles.

As with the halo analysis, it was important to establish if the simulations had adequately reproduced known relationships. Therefore, the relationship between metallicity and galactocentric radius was investigated.

Each galaxy was then analysed in isolation to determine if solar type particles in Milky Way type galaxies have unique evolutionary histories. The solar type particles from each galaxy were then compared with each other.

For both the isolated and inter-galaxy analysis, the galactocentric radius, mass-weighted time, expected number of supernova experienced in the first 100 Myrs since formation, and the amount of time spent in a cluster environment during the first 100 Myrs since formation were utilised. The following section describes the processes used and the results obtained.

5.1 Using the Metallicity of Galaxies to Identify Major Mergers

To establish if the galaxies replicate the metallicity gradient seen in observations, the mean galactocentric radius was taken for each group in each galaxy. The results are outlined in Table 5.1, and illustrative examples can be found in Figures 5.1, 5.2, and 5.3. The errors on the values in Table 5.1 are of the order of 10^{-3} - 10^{-4}, meaning they do not make a difference to the conclusions that have been drawn, and they are not included here.

Figures 5.1, 5.2, and 5.3 all show metal poor particles for any given metallicity having larger fractional radii than particles with higher metallicities. Metal rich particles are shown to be found nearer the centre of their host galaxy, with solar metallicity particles being found between them. It should be noted that there are still significant overlaps in the distributions of the different metallicity groups, meaning that particles of different metallicities are not segregated within their galaxy.

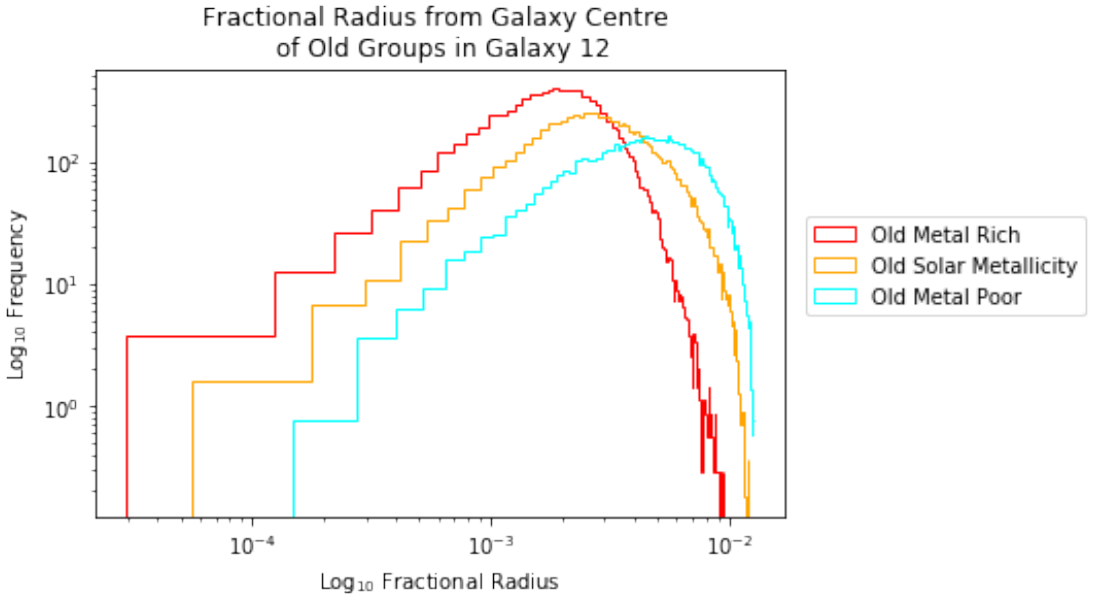

FIGURE 5.1: Fractional distance distribution of old particles in galaxy 12

As the distributions of the particles are not symmetrical, the mean values can be significantly skewed by outliers. In order to provide another comparison that is not as seriously impacted by outliers, the median values were also calculated. These can be found in Table 5.2.

It is expected that higher metallicity particles will be found at lower galactocentric radii, with metal poor particles being found at the largest radii. This is found to be the case in both the mean and median for 7 galaxies (6, 10, 11, 12, 14, 20, and 24).

TABLE 5.1: Mean values for the galactocentric radius of each group in each galaxy. Results in italics represent a group with less than 100 particles.

Galaxy Number	Mean Relative Galactocentric Radius (%)							
	Old Rich	Solar Rich	Young Rich	Old Solar	Solar	Young Solar	Old Poor	Solar Poor	Young Poor
1	0.248	0.215	0.193	0.302	0.317	0.358	0.347	n/a	*0.378*
2	0.203	0.187	0.153	0.297	*0.394*	0.402	0.395	*0.499*	0.515
3	*0.607*	0.481	0.386	0.733	0.575	0.902	0.931	n/a	n/a
4	*0.380*	*0.589*	0.270	0.698	0.630	0.781	1.184	*2.405*	*2.089*
5	0.595	0.553	0.400	0.688	0.794	0.928	0.916	n/a	n/a
6	*0.755*	*0.739*	*0.696*	1.236	1.373	1.464	1.723	*3.907*	*3.061*
10	0.175	0.148	0.134	0.219	0.227	0.301	0.305	*0.375*	0.384
11	n/a	n/a	*0.274*	0.355	*0.332*	0.359	0.578	0.474	0.542
12	0.236	0.208	0.187	0.371	0.337	0.498	0.546	0.742	0.745
13	0.348	0.292	0.254	0.556	0.741	0.712	0.686	*1.044*	1.017
14	*0.117*	n/a	0.054	0.169	0.197	0.175	0.245	*0.337*	0.370
15	n/a	*1.019*	0.758	1.426	*1.458*	1.133	1.561	1.410	1.458
20	n/a	n/a	n/a	0.548	0.473	0.507	1.047	0.990	1.231
22	0.457	0.913	0.203	0.305	0.329	0.281	0.519	0.682	0.646
23	0.332	0.265	0.180	0.445	0.539	0.620	0.517	n/a	n/a
24	*0.204*	*0.293*	0.170	0.392	0.371	0.381	0.565	*0.866*	0.891

TABLE 5.2: Scaled median values for the galactocentric radius of each group in each galaxy. Results in italics represent a group with less than 100 particles.

Galaxy Number	Median Relative Galactocentric Radius (%)							
	Old Rich	Solar Rich	Young Rich	Old Solar	Solar	Young Solar	Old Poor	Solar Poor	Young Poor
1	0.249	0.199	0.178	0.289	0.302	0.354	0.341	n/a	*0.378*
2	0.194	0.180	0.140	0.282	*0.407*	0.400	0.531	*0.499*	0.531
3	*0.511*	0.392	0.286	0.688	0.472	0.910	0.948	n/a	n/a
4	*0.318*	*0.539*	0.216	0.585	0.498	0.592	1.030	*2.495*	*2.162*
5	0.556	0.504	0.301	0.676	0.805	0.950	0.942	n/a	n/a
6	*0.734*	*0.772*	*0.442*	1.104	1.226	1.247	1.587	*3.907*	*2.968*
10	0.167	0.144	0.129	0.208	0.210	0.295	0.296	*0.374*	0.387
11	n/a	n/a	*0.240*	0.309	*0.279*	0.306	0.540	0.420	0.508
12	0.221	0.197	0.172	0.333	0.283	0.474	0.532	0.763	0.756
13	0.323	0.273	0.236	0.516	0.730	0.693	0.659	*1.016*	1.017
14	*0.110*	n/a	0.046	0.148	0.176	0.143	0.227	*0.331*	0.374
15	n/a	*1.019*	0.629	1.373	*1.336*	1.009	1.518	1.326	1.416
20	n/a	n/a	n/a	0.460	0.392	0.419	0.889	0.837	1.121
22	0.220	0.958	0.105	0.216	0.234	0.188	0.479	0.690	0.652
23	0.275	0.224	0.148	0.406	0.532	0.617	0.510	n/a	n/a
24	*0.149*	*0.279*	0.140	0.453	0.311	0.318	0.536	*0.933*	0.884

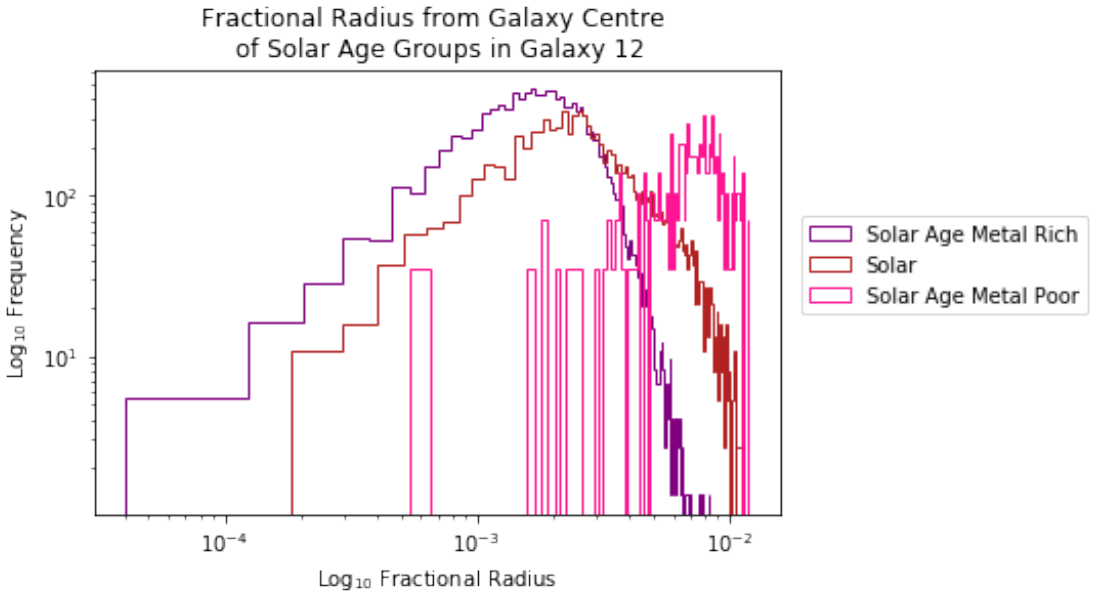

FIGURE 5.2: Fractional distance distribution of solar age particles in galaxy 12

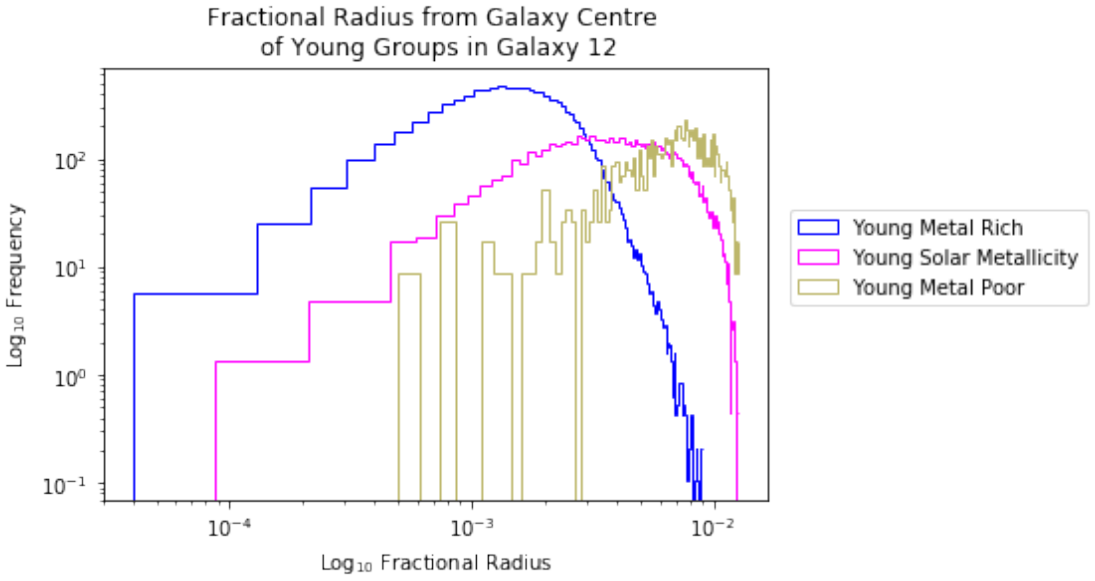

FIGURE 5.3: Fractional distance distribution of young particles in galaxy 12

In galaxy 1, young solar metallicity particles have higher mean and median values than old metal poor particles, but these differences are small, with the mean being higher by 0.011%, and the median being higher by 0.013%. Neither of these groups contains less than 100 particles, and there is no known merger in this galaxy, so the reason for this divergence from the expected trend is unknown.

In galaxies 2 and 5, the medians are found to follow the expected trend. However, when the means are examined, the young solar metallicity particles in both galaxies have higher galactocentric radii than old metal poor particles. However, these differences are

small at 0.007% and 0.012% respectively. Whilst there are no obvious reasons for this in galaxy 5, galaxy 2 only has 518 young solar metallicity particles, which may limit the accuracy of the measurement for this group.

In galaxy 3, old metal rich particles have a higher mean and median galactocentric radius than solar particles. However, the difference in these values is only 0.032% and 0.039% respectively. In addition to this, the old metal rich group in this galaxy contains only 55 particles, so these results may have been skewed by the small sample size. On further inspection of this galaxy, the old metal rich particles were found to start at a galactocentric radius of 0.231kpc, which is much larger than any other group. Whilst there is no obvious cause for this, it has the effect of increasing the mean galactocentric radius.

Galaxy 4 has mean values which follow the proposed trend, but the median value for solar age metal rich particles is higher than that for the solar group. However, the difference is only 0.041%, and the solar age metal rich group has less than 100 particles, meaning the small sample size may have impacted the median.

In galaxy 13, both solar and young solar particles have higher means and medians than the old metal poor particles. The solar group has a mean 0.055% larger than that of the old metal poor group, and a median 0.071% larger. For the young solar metallicity group, these values are 0.026% and 0.034% respectively. None of these groups contain less than 100 particles, so there is no obvious reason for this to have occurred.

Galaxy 15 shows old solar metallicity particles having higher mean galactocentric radii than solar age metal poor particles. However, the difference is only 0.031 kpc. Solar particles have a higher mean galactocentric radius than both solar age metal poor and young metal poor particles. However, the differences are 0.094 kpc and 0.001 kpc respectively.

In galaxy 15, the solar poor group has smaller mean and median values than both the old solar metallicity and solar metallicity groups. For the old solar metallicity group, the difference in the mean is 0.016% and the difference in the median in 0.047%. For the solar metallicity group, these values are 0.048% and 0.010% respectively. Whilst there is no explanation for these differences in the old solar group, the solar group contains less than 100 particles, which can skew values obtained by statistical tests.

The young solar metallicity group in galaxy 15 has a slightly lower median value than the solar age metal rich group. The difference is small, at only 0.010%, as is likely explained by the lack of particles in the solar age metal rich group.

In galaxy 22, solar age metal rich particles have the highest mean and median galactocentric radius. The differences in the mean of this group and those of the solar metallaicity and metal poor particles is between 0.231% and 0.632%. For the median values, this range is between 0.268% and 0.770%.

Also in galaxy 22, old metal rich particles have a mean galactocentric radius higher than all solar metallicity particles, and a median higher than old and young solar metallicity particles.

These results may be due to some of the previously identified merger particles having migrated into the central galaxy and skewed the results. It is also possible that the merger caused some particles to be perturbed, moving them from their original positions.

In galaxy 23, old metal poor particles have a lower mean and median galactocentric radius than both solar and young solar metallicity particles. The differences in the mean values are 0.022% and 0.103% respectively, and the differences in the median values are 0.022% and 0.107%.

Whilst the deviations from the known trend are unexpected, some can be put down to small group sizes and outliers impacting the accuracy of the calculations. It is also notable that many of the deviations identified are less than 0.1% of the size of the galaxy in question, suggesting that these differences are insignificant.

5.2 Comparisons Within Galaxies

5.2.1 Statistical Analysis of the Galactocentric Radius

As with the analysis carried out within each halo, the distributions of the galactocentric radii of the particles were compared to the distribution of the solar group. These comparisons were done with both Kolmogorov-Smirnov and Anderson-Darling tests, and the same 2σ cut off was applied.

The results found to be above the 2σ cut off are outlined in Table 5.3. Results in italics signify results that are of limited statistical use due to small sample sizes.

Many of the Anderson-Darling tests give the same $\log_{10}$ p-value due to the limits imposed on the p-values. The upper limit, which suggests a high level of similarity, is 0.25, which gives a $\log_{10}$ p-value of 0.602. This means that once a certain level of similarity has been reached, all p-values will be the same.

As the Kolmogorov-Smirnov and Anderson-Darling test both compare the whole distribution, any groups with only one value above the cut off are discarded, as distributions

TABLE 5.3: Statistically significant results of Kolmogorov-Smirnov and Anderson-Darling tests comparing the galactocentric radius within galaxies

Galaxy	Group	$\log_{10}$ KS p-value	$\log_{10}$ AD p-value
1	*Young Metal Poor*	*-0.290*	
1	*Solar Age Metal Rich*		*-0.602*
2	*Old Metal Poor*	*-0.192*	*-0.602*
2	*Solar Age Metal Poor*	*-0.525*	*-0.602*
2	*Young Solar Metallicity*	*-0.039*	*-0.602*
3	*Old Metal Rich*	*-0.231*	*-0.602*
4	*Solar Age Metal Rich*	*-0.374*	*-0.602*
4	*Solar Age Metal Poor*	*-1.006*	
6	*Solar Age Metal Rich*		*-1.276*
6	*Solar Age Metal Poor*	*-1.010*	
6	Young Solar Metallicity	-0.589	-0.696
11	*Old Solar Metallicity*	*-0.414*	*-0.602*
11	*Young Metal Rich*	*-0.162*	*-0.602*
11	*Young Solar Metallicity*	*-0.488*	*-0.602*
15	*Old Solar Metallicity*	*-0.089*	*-0.602*
15	*Old Metal Poor*	*-0.543*	*-0.602*
15	*Solar Age Metal Rich*	*-0.230*	*-0.602*
15	*Solar Age Metal Poor*	*-0.063*	*-0.602*
15	*Young Metal Poor*	*-0.098*	*-0.602*
20	Young Solar Metallicity	-0.831	-0.747
24	*Solar Age Metal Rich*	*-0.313*	*-0.602*
24	Young Solar Metallicity	-0.784	

with significant similarities should give a value for each test. This means both results from galaxy 1, the solar age metal poor result from galaxies 4 and 6, the solar age metal rich result from galaxy 6, and the young solar metallicity result from galaxy 24 are discarded.

This left 16 pairs of results, 14 of which are limited by small sample sizes. The two pairs of results that are not limited by sample size are young solar metallicity groups, and 9 of the 14 other pairs are either solar age or metallicity. This suggests that groups with solar properties have more similarities to their respective solar groups than groups without solar properties. However, this is not seen across all galaxies, so it cannot be said that this trend is consistent throughout the simulations.

These results show the strongest similarities between groups that have similar mean and median galactocentric radii (see Section 5.1), meaning the statistical tests support the previous conclusions.

Overall, these tests do not show a strong similarity between the solar distribution and that of other groups. Whilst particles with a solar property were more likely to have a $\log_{10}$ p-value above the cut off, the vast majority of these were limited due to sample

sizes, making them less useful for establishing trends. It is also notable that there are no groups that produce results above the 2σ cut off in each galaxy. This suggests that either the solar distributions vary between galaxies, the distribution within each group vary between galaxies, or both distributions vary between galaxies. Whichever of these is true, this suggests that each galaxy has its own unique populations in each age-metallicity group, meaning the formation and evolution of these particles is likely due to the environments in which they reside.

5.2.2 Statistical Analysis of the Mass Weighted Time

Any results found to be above the 2σ cut off from the analysis of the mass-weighted time are detailed in Table 5.4.

TABLE 5.4: Results from both the Kolmogorov-Smirnov and Anderson-Darling tests used to analyse mass-weighted time within galaxies. Results in italics indicate a small sample size.

Galaxy	Group	$\log_{10}$ KS p-Value	$\log_{10}$ AD p-Value
1	*Young Metal Poor*	*-0.760*	*-0.671*
2	*Old Solar Metallicity*	*-0.739*	*-1.032*
2	*Solar Age Metal Poor*	*-0.987*	
2	*Young Solar Metallicity*		*-1.268*
4	*Old Metal Rich*		*-1.153*
4	*Solar Age Metal Rich*	*-0.368*	*-0.602*
4	*Solar Age Metal Poor*	*-0.660*	*-0.602*
6	*Old Metal Rich*	*-0.256*	*-0.602*
6	*Solar Age Metal Rich*	*-0.605*	*-0.602*
6	*Solar Age Metal Poor*	*-0.216*	*-0.602*
6	*Young Metal Rich*	*-0.180*	*-0.602*
11	*Solar Age Metal Poor*	*-0.596*	*-0.602*
11	*Young Metal Rich*	*-1.239*	*-0.602*
11	*Young Solar Metallicity*	*-0.759*	*-0.906*
11	*Young Metal Poor*	*-0.192*	
14	*Old Metal Rich*	*-0.328*	*-0.729*
15	*Old Solar Metallicity*	*-0.052*	*-0.602*
15	*Old Metal Poor*	*-0.118*	*-0.602*
15	*Solar Age Metal Rich*	*-0.081*	*-0.602*
15	*Solar Age Metal Poor*	*-1.089*	
20	Old Solar Metallicity	-0.979	
22	Old Metal Rich	-1.069	-1.148
24	*Old Metal Rich*	*-0.899*	*-0.891*
24	*Solar Age Metal Rich*	*-1.043*	*-0.754*

As outlined in Section 5.2.1, there are upper and lower limits on the values produced by the Anderson-Darling test, meaning many of the results are the same.

As each test supports the other, groups where only one of the tests has produced a $\log_{10}$ p-value above the 2σ cut off are discarded and not considered when establishing trends.

This leaves 18 pairs of results, 17 of which are limited due to small sample sizes. Of the 18 pairs of results, half of them include groups with solar ages or metallicities, which may suggest that groups with these properties are more likely to display similarities to their corresponding solar group.

As the time intervals used for calculating MWT are often the same for each particle, having a similar MWT suggests that the particles lost their mass at a similar rate. As half of the pairs with results above 2σ include groups with solar ages or metallicities, this result implies that these groups lose their cluster mass at a similar rate to the solar group within their galaxy.

5.2.3 Statistical Analysis of the Number of Supernova

The number of supernova experienced in the first 100 Myrs since formation was also analysed using Kolmogorov-Smirnov and Anderson- Darling tests. These results are shown in Table 5.5.

As with previous tests, for results to be considered robust the group must have both a Kolmogorov-Smirnov and Anderson-Darling result above the 2σ cut off. Applying this to Table 5.5 leaves 30 pairs of results. Six of these pairs are not limited by small sample sizes, and five of these six pairs include a group with a solar type property. This may suggest that groups with solar type properties have more similar distributions to their respective solar groups than those without solar properties. This is supported by fifteen additional pairs of results that also have solar properties, although these are limited by small sample sizes.

5.2.4 Statistical Analysis of the Time Spent in a Cluster (0-100 Myrs)

Finally, the amount of time particles spent in a cluster environment was compared. These results are found in Table 5.6.

Four results from Table 5.6 were discounted as they lacked a Kolmogorov-Smirnov test result above the 2σ cut off. This left 23 pairs, all of which were limited by small sample sizes. This means that any conclusions drawn from these statistical tests would need to be confirmed through other channels (e.g. other simulations).

Of the 23 pairs, thirteen involve groups with solar age or metallicity. This may suggest that these groups show more similarities to their corresponding solar distribution than

TABLE 5.5: Statistical analysis of the distribution of the expected number of supernovae experienced in the first 100 Myrs since formation compated to solar groups.

Galaxy	Group	KS $\log_{10}$ p-Value	AD Test $\log_{10}$ p-Value
1	Old Solar Metallicity	-0.762	-0.636
1	*Young Metal Poor*	*-1.013*	
2	*Old Solar Metallicity*	*-0.758*	*-0.778*
2	*Old Metal Poor*	*-0.230*	*-0.602*
2	*Solar Age Metal Poor*	*-0.914*	*-1.169*
2	*Young Solar Metallicity*	*-0.513*	*-0.877*
2	*Young Metal Poor*	*-0.848*	*-1.286*
3	*Old Metal Rich*	*-0.120*	*-0.602*
4	*Old Metal Rich*	*-0.334*	*-0.602*
4	*Solar Age Metal Rich*	*-0.067*	*-0.602*
4	*Solar Age Metal Poor*	*-0.885*	*-1.004*
6	*Old Metal Rich*	*-0.607*	*-0.602*
6	Old Solar Metallicity	-0.141	-0.602
6	*Solar Age Metal Poor*	*-0.335*	*-0.602*
6	*Young Metal Rich*	*-0.040*	*-0.602*
6	Young Solar Metallicity	-1.089	-1.121
11	*Old Solar Metallicity*	*-0.523*	*-0.981*
11	*Solar Age Metal Poor*	*-0.151*	*-0.602*
11	*Young Metal Rich*	*-1.131*	*-0.741*
11	*Young Solar Metallicity*	*-0.039*	*-0.602*
11	*Young Metal Poor*	*-0.713*	*-1.253*
13	Young Solar Metallicity	-0.043	-0.602
14	*Old Metal Rich*	*-0.124*	*-0.602*
14	Old Solar Metallicity	-0.062	-0.602
15	*Old Solar Metallicity*	*-0.540*	*-0.602*
15	*Old Metal Poor*	*-0.788*	*-0.602*
15	*Solar Age Metal Rich*	*-0.440*	*-0.602*
15	*Solar Age Metal Poor*		*-1.004*
15	*Young Metal Poor*	*-1.197*	*-0.631*
22	Old Metal Rich	-0.705	-0.816
24	*Old Metal Rich*	*-0.573*	*-0.602*
24	*Solar Age Metal Rich*	*-0.016*	*-0.602*

groups without a solar property, but the limited sample sizes mean this cannot be confirmed. It should also be noted that these results only include six of the 16 galaxies, and when all the results are considered, only eight galaxies are shown. This suggests that this trend is not seen across all galaxies, and that this may be due to the initial conditions used in these particular galaxies.

5.2.5 Conclusions

Each of the properties tested showed that groups with either a solar age or metallicity were more likely to have $\log_{10}$ p-values above the 2σ cut off than groups without a solar

TABLE 5.6: Results from the Kolmogorov-Smirnov and Anderson-Darling tests comparing the amount of time spent in a cluster environment in the first 100 Myrs since formation.

Galaxy	Group	KS $\log_{10}$ p-Value	AD Test $\log_{10}$ p-Value
1	*Young Metal Poor*	*-0.902*	*-1.069*
2	*Old Metal Rich*	*-1.068*	*-1.007*
2	*Old Solar Metallicity*	*-0.839*	*-0.807*
2	*Old Metal Poor*	*-1.292*	*-1.231*
2	*Solar Age Metal Poor*	*-0.181*	*-0.602*
2	*Young Solar Metallicity*	*-0.612*	*-0.966*
4	*Old Metal Rich*	*-0.557*	*-0.602*
4	*Solar Age Metal Rich*	*-0.587*	*-0.759*
4	*Solar Age Metal Poor*	*-0.838*	*-0.851*
6	*Old Metal Rich*	*-0.299*	*-0.602*
6	*Solar Age Metal Rich*		*-1.027*
6	*Solar Age Metal Poor*	*-0.099*	*-0.602*
6	*Young Metal Rich*	*-1.280*	*-0.932*
11	*Old Solar Metallicity*		*-1.098*
11	*Old Metal Poor*	*-1.131*	*-0.747*
11	*Solar Age Metal Poor*	*-0.302*	*-0.602*
11	*Young Metal Rich*	*-0.003*	*-0.602*
11	*Young Solar Metallicity*	*-0.573*	*-0.602*
11	*Young Metal Poor*		*-1.020*
14	*Old Metal Rich*	*-0.231*	*-0.602*
15	*Old Solar Metallicity*	*-0.644*	*-0.602*
15	*Old Metal Poor*	*-0.274*	*-0.602*
15	*Solar Age Metal Rich*	*-0.641*	*-0.602*
15	*Solar Age Metal Poor*	*-0.423*	*-0.685*
15	*Young Metal Rich*		*-1.128*
15	*Young Solar Metallicity*	*-0.109*	*-0.602*
24	*Solar Age Metal Rich*	*-0.580*	*-0.602*

property. However, many of the results are limited by small sample sizes, which can impact the accuracy of both tests, so these results cannot be considered conclusive. It was also apparent that this trend was not shown by all galaxies as galaxies 3, 5, 10, 12, and 23 did not produce any pairs of results above the 2σ cut off for any of the statistical test. This may suggest that these trends are dependent on the evolutionary histories that particles have, which themselves may be impacted by the initial conditions.

Many of the results obtained from the statistical tests were below the 2σ cut off, suggesting that there is limited similarity between the solar particles and the particles in other distributions in most cases. This suggests that similarities in distributions are rare, and do not occur consistently between galaxies.

5.3 Statistical Comparison Between Galaxies

As it has now been shown that solar type particles have limited similarities in their evolutionary histories when compared to other particles in their respective galaxies, it is important to establish if solar particles from different galaxies share similar evolutionary histories. To do this, Kolmogorov-Smirnov and Anderson Darling tests were carried out on the galactocentric radius, mass-weighted time, the expected number of supernova in the first 100 Myrs, and the amount of time spent in a cluster in the first 100 Myrs.

5.3.1 Comparison of the Galactocentric Radii Between Galaxies

As Milky-Way type galaxies are defined only by their masses in the EMP simulations, the galaxies produced vary in radius. Therefore, in order to compare the distributions of solar type particles between galaxies, the distributions were scaled. This was done by dividing the calculated galactocentric radii of the particles by the maximum galactocentric radii in the galaxy. Both the Kolmogorov-Smirnov and Anderson-Darling tests were then carried out on these scaled distributions.

TABLE 5.7: Kolmogorov-Smirnov and Anderson-Darling results comparing the distributions of the galactocentric radii of solar groups from different galaxies.

Galaxy Numbers	KS $\log_{10}$ p-Value	AD $\log_{10}$ p-Value
11 and 1	*-0.314*	*-1.024*
11 and 12	*-0.016*	*-0.602*
15 and 6	*-0.079*	*-0.602*
11 and 22		*-1.065*
11 and 24	*-0.611*	*-0.602*

As the comparison between galaxies 11 and 22 does not include a Kolmogorov-Smirnov result, this pair is not considered when determining if there are any trends. This leaves 4 pairs, all of which are limited by small sample sizes. Galaxy 11 appears in 3 of the 4 remaining pairs, which suggests that the galactocentric distribution in this galaxy is the most similar to that found in other galaxies. However, the limited sample size may have impacted the statistical results.

As the statistical tests produced 120 pairs of values but only four are above the 2σ cut off, it is reasonable to state that there is very little similarity between the shapes of most solar distributions from different galaxies. This suggests that initial and environmental conditions are important in determining where solar type particles are found in their host galaxy, and this is not a fundamental property of the particles themselves.

5.3.2 Comparison of the Mass-Weighted Time Between Galaxies

The same tests were employed to analyse the mass-weighted time between galaxies, and the results are shown in Table 5.8.

TABLE 5.8: Statistical Analysis of the mass-weighted time distributions of solar particles from different galaxies.

Galaxy Numbers	KS $\log_{10}$ p-Value	AD $\log_{10}$ p-Value
1 and *2*	*-0.123*	
1 and *11*	*-1.080*	
2 and 6	*-1.237*	
2 and *11*	*-0.828*	
2 and 12	*-1.083*	
2 and 13	*-1.130*	
2 and *15*	*-1.235*	
2 and 20	*-1.002*	
4 and *15*	*-0.404*	
6 and *11*	*-0.960*	
6 and *15*	*-0.159*	
6 and 20	-0.756	
11 and 4		*-1.288*
11 and 14		*-0.602*
11 and *15*	*-0.579*	
11 and 20	*-1.242*	*-0.602*
13 and *15*	*-0.372*	
15 and 20	*-0.436*	
15 and 22	*-0.493*	
15 and 24	*-0.545*	
20 and 24	-1.125	

Only one pair of values above the 2σ cut off is produced when the mass-weighted time is analysed. The remaining pair is limited by a small sample size, meaning it is not possible to draw any conclusions from this analysis.

5.3.3 Comparison of the Expected Number of Supernova Between Galaxies

The distribution of the expected number of supernova in each pair of galaxies was also compared. This gave the results in Table 5.9.

This analysis produces 14 pairs of results, only four of which are not limited by a small sample size. These four results are between galaxies 4 and 24, 6 and 14, 6 and 20, and 14 and 20. Three of these results include galaxies 6, 14, and 20, which suggests that these galaxies all have similar supernovae distributions. When the values from

TABLE 5.9: Statistical Analysis of the expected number of supernova between galaxies

Galaxy Numbers	KS $\log_{10}$ p-Value	AD $\log_{10}$ p-Value
1 and *2*	*-0.667*	*-1.221*
2 and 4	*-1.000*	
2 and 13	*-0.544*	*-0.742*
2 and 24	*-0.800*	
4 and 24	-0.619	-0.671
6 and *11*	*-0.037*	*-0.602*
6 and 14	-0.450	-0.602
6 and *15*	*-0.449*	*-0.602*
6 and 20	-0.555	-0.770
11 and 14	*-0.045*	*-0.602*
11 and *15*	*-0.272*	*-0.602*
11 and 20	*-0.068*	*-0.602*
14 and *15*	*-0.817*	*-0.602*
14 and 20	-0.163	-0.602
15 and 20	*-0.894*	*-0.602*
15 and 22	*-1.020*	*-1.269*

these results are examined, the Kolmogorov-Smirnov results are all above -0.6, and the Anderson-Darling results are all above -0.8, confirming the similarity.

If these 3 galaxies are examined further, it can be seen that they all have significant Kolmogorov-Smirnov and Anderson-Darling results with galaxies 11 and 15. Whilst it should be noted that the solar groups in these galaxies contain less than 100 particles, when galaxy 11 is compared to galaxies 6, 14, and 20, all the Kolmogorov- Smirnov values are above -0.07 and all the Anderson-Darling values are -0.602, which indicates the highest level of similarity. When galaxies 6, 14, and 20 are compared to galaxy 15, the Kolmogorov- Smirnov results are all above -0.9, and all the Anderson-Darling results are -0.602, again indicating the higest level of similarity. When galaxies 11 and 15 are compared to each other, the Anderson- Darling result is again -0.602, and the Kolmogorov-Smirnov result is -0.272, both of which suggest high levels of similarity.

The results from these 5 galaxies account for 10 of the 14 pairs of results, suggesting that distributions across these 5 galaxies are similar to each other. In addition to this, nine of the ten Anderson- Darling results have the highest level of similarity, further supporting the theory that these galaxies have similar distributions of supernovae events in the first 100 Myrs.

5.3.4 Comparison of the Amount of Time Spent in a Cluster Environment Between Galaxies

Finally, the amount of time solar particles spent in a cluster environment was investigated. The results are shown in Table 5.10.

TABLE 5.10: Statistical Analysis of the amount if time spent in a cluster environment between galaxies

Galaxy Numbers	KS $\log_{10}$ p-Value	AD $\log_{10}$ p-Value
1 and *2*	*-0.289*	*-0.602*
2 and 4	*-0.439*	*-0.602*
2 and 13	*-0.232*	*-0.602*
2 and 24	*-0.316*	*-0.602*
6 and *15*	*-0.138*	*-0.602*
11 and *15*	*-1.013*	*-0.816*
11 and 22	*-0.248*	*-0.602*
15 and 20	*-0.745*	*-1.031*

This gave 8 pairs of results above the 2σ cut off, which suggests that certain pairs of galaxies have similar distributions in the amount of time solar particles spend in a cluster environment. However, this is not repeated across all possible galaxy combinations, suggesting that solar particles do not spend the same amount of time in cluster environments in each galaxy. It is also important to note that all of the results are compromised due to small sample sizes.

5.3.5 Conclusions

The results of the statistical tests do not suggest that there is a universal underlying distribution that all solar particles are drawn from. However, many of the results were limited by small sample sizes, and only solar type stars were compared, so it is possible that other groups may share an underlying distribution across galaxies.

The results for the number of supernova experienced in the first 100 Myrs did show that some galaxies experienced similar distributions of supernova in their early evolutionary histories. Galaxies 6, 11, 14, 15, and 20 all had similar distributions to each other. Taking the particles in these galaxies to be proxies for solar type stars, this suggests that solar type stars in these galaxies would potentially be subjected to similar distributions of supernovae in their early lives.

As supernovae can impact the formation and evolution of planets, this result would benefit from future examination. Ideally, this would utilise simulations with the resolution

needed to model individual stars and planets, as this would enable additional properties, such as the distance between stars, planets, and supernova events, to be modelled.

No results could be drawn from the analysis of the mass-weighted time due to a lack of statistically meaningful samples, and the results from the analysis into both the galactocentric radius and the amount of time spent in a cluster environment in the first 100 Myrs are limited due to small sample sizes.

5.4 Establishing Links Between Metallicity and Evolutionary History

As with the halo analysis conducted in Chapter 4, it has now been established that solar type particles have different evolutionary histories to other particles types. As metallicity-dependent trends have previously been identified, this raises the question of whether the evolutionary history is determined by metallicity. To this end, the initial fraction of mass in a cluster environment, the cluster disruption time, the amount of supernova experienced in the first 100 Myrs since formation, and the amount of time spent in a cluster environment in the first 100 Myrs since formation have been analysed with respect to the metallicity of the particles.

5.4.1 Link Between the Initial Fraction of Mass in a Cluster Environment and Metallicity

As with the halo analysis in Section 4.4.1, it was noted that there was a correlation between the initial fraction of mass in a cluster environment and the metallicity of the particle. Therefore, this value was recorded for each group in each galaxy, along with the standard error on the mean, and these results are outlined in Table 5.11. Examples of the evolution of the mass fraction in a cluster environment can be found in Figures 5.4 and 5.5.

Both Figures 5.4 and 5.5 illustrate that metal rich particles (red) have the highest initial fraction of mass in a cluster environment, followed by solar metallicity particles (cyan and magenta), and then metal poor particles (blue). On initial inspection, Figure 5.4 appeared to suggest an exponential decrease in the fraction of mass as the particles evolved over time. However, closer inspection via a logged plot (Figure 5.5) revealed this was not the case. It became evident that mass is lost very quickly in the early lives of all particles, with this rate slowing down as the particles evolve, likely due to a reduction in tidal disruption, which in turn decreases the rate of mass loss. This

TABLE 5.11: The mean initial percentage of mass in a cluster environment for all groups in all galaxies.

Galaxy Number	Mean Percentage of Initial Mass in Cluster Environment							
	Old Rich	Solar Rich	Young Rich	Old Solar	Solar	Young Solar	Old Poor	Solar Poor	Young Poor
1	39.7 ±1.1	38.7 ±0.9	41.2 ±0.1	33.6 ±0.1	32.9 ±0.3	31.1 ±0.2	31.2 ±0.1	n/a	*0.0* ±0.0
2	48.8 ±0.1	50.8 ±0.2	58.9 ±0.1	41.2 ±0.1	*34.1*±3.2	31.9 ±0.8	34.3 ±0.1	*11.8* ±0.0	28.5±0.5
3	*44.0* ±1.9	45.0 ±0.5	52.2 ±0.0	46.5 ±0.2	45.9 ±0.3	47.6 ±0.2	47.3 ±0.3	n/a	n/a
4	*29.0* ±1.5	*30.1* ±5.8	37.4 ±0.5	22.9 ±0.1	24.8 ±0.3	23.8 ±0.1	18.1 ±0.1	*6.5* ±0.0	*14.3* ±1.4
5	57.9 ±0.1	58.1 ±0.2	60.4 ±0.0	57.0 ±0.2	58.3 ±0.5	56.5 ±0.6	57.2 ±0.6	n/a	n/a
6	*20.1* ±2.2	*26.7* ±4.7	*20.2* ±3.2	19.9 ±0.2	18.7 ±0.7	18.0 ±0.2	17.9 ±0.1	*21.9* ±0.0	*11.4* ±1.0
10	52.8 ±0.1	56.0 ±0.1	58.4 ±0.0	48.3 ±0.1	47.6 ±0.3	40.5 ±0.2	40.0 ±0.1	*33.1* ±2.0	34.2 ±1.3
11	n/a	n/a	*14.1* ±0.9	18.5 ±0.2	*23.8* ±3.4	19.9 ±0.1	16.2 ±0.1	18.1 ±0.3	17.1 ±0.2
12	46.9 ±0.1	52.7 ±0.1	56.5 ±0.0	36.3 ±0.1	40.9 ±0.3	31.2 ±0.1	28.0 ±0.1	20.1 ±0.9	21.4 ±0.5
13	46.4 ±0.1	52.7 ±0.2	57.4 ±0.0	35.5 ±0.1	29.1 ±0.6	30.2 ±0.1	31.0 ±0.1	*22.0* ±2.2	22.6 ±0.3
14	*22.0* ±3.3	n/a	39.5 ±0.1	21.4 ±0.1	20.9 ±0.3	23.9 ±0.1	19.4 ±0.1	*15.4* ±1.5	17.2 ±0.2
15	n/a	*26.5* ±0.0	31.1 ±1.0	20.5 ±0.6	*20.1* ±2.2	24.7 ±0.1	21.4 ±0.2	22.2 ±0.3	21.9 ±0.2
20	n/a	n/a	n/a	18.2 ±0.2	19.7 ±0.7	18.0 ±0.2	15.4 ±0.1	15.7 ±0.4	14.7 ±0.1
22	15.2 ±0.8	11.7 ±0.7	30.7 ±0.1	16.4 ±0.1	15.9 ±0.2	18.4 ±0.0	14.8 ±0.1	12.5 ±0.6	12.0 ±0.3
23	54.9 ±0.1	56.9 ±0.2	61.3 ±0.0	52.9 ±0.1	53.8 ±0.8	53.0 ±0.6	52.3 ±0.4	n/a	n/a
24	*22.3* ±1.8	*24.7* ±4.3	40.1 ±0.1	24.3 ±0.1	25.9 ±0.2	27.0 ±0.1	21.0 ±0.1	*10.2* ±1.7	17.0 ±1.2

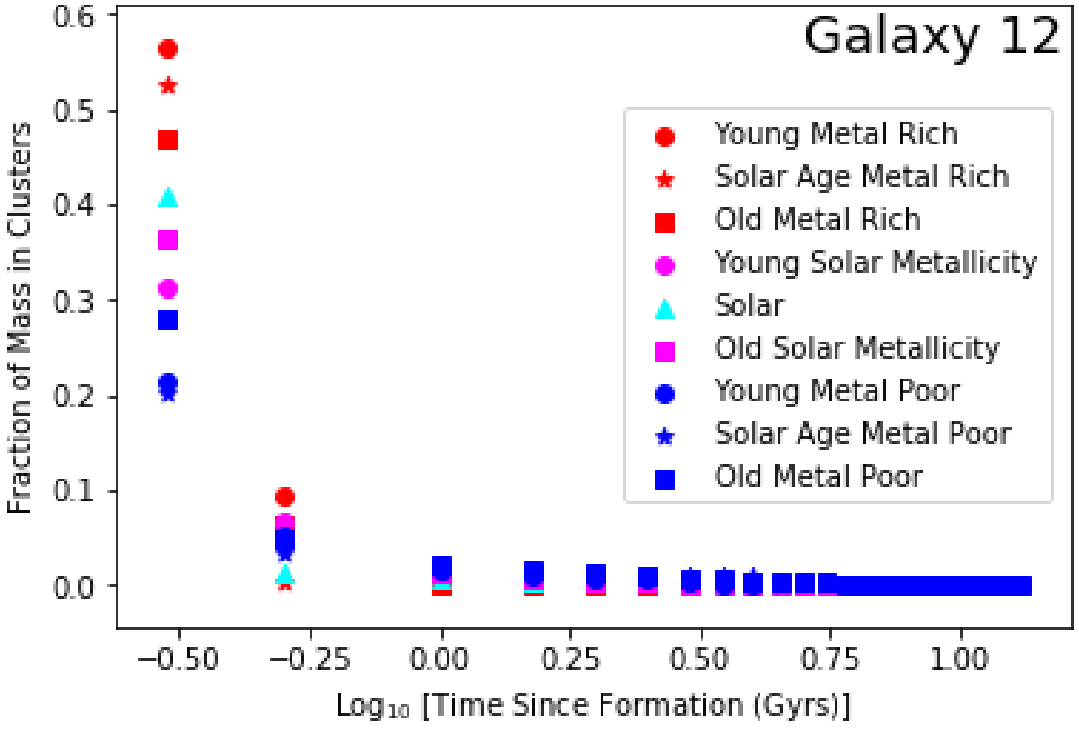

FIGURE 5.4: Initial fraction of mass in a cluster environment for all age-metallicity groups in galaxy 12.

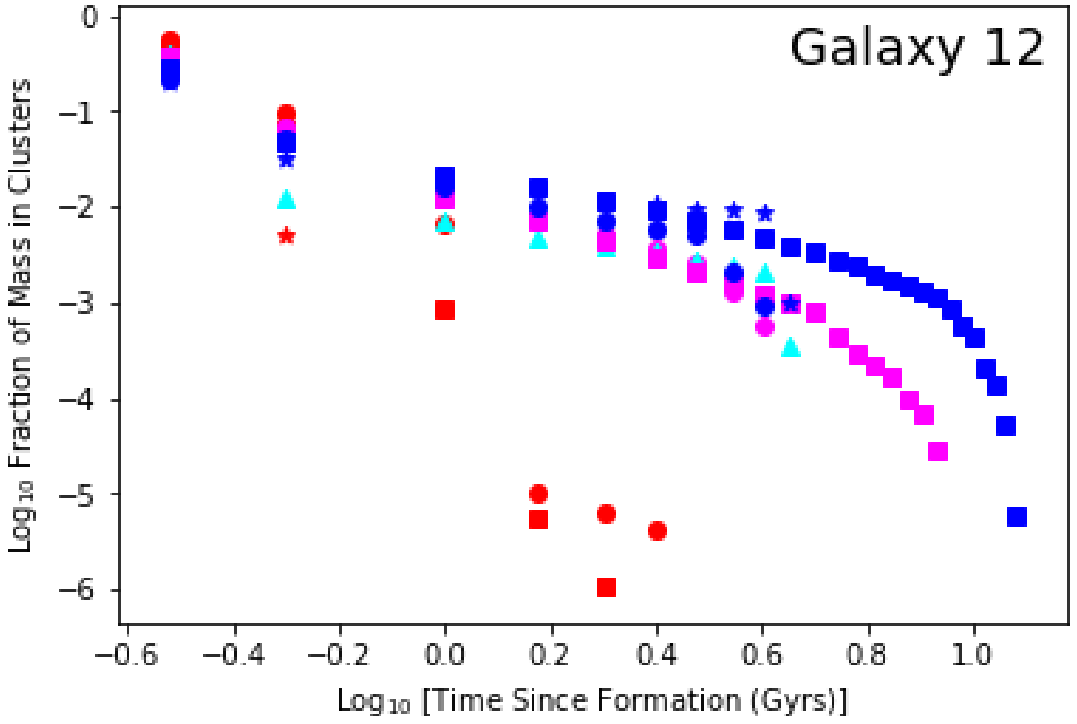

FIGURE 5.5: Initial fraction of mass in a cluster environment for all age-metallicity groups in galaxy 12 on logged axis.

process then accelerates again at later times. Figure 5.5 also shows a distinct difference in the evolution of the fraction of mass in a cluster environment for metal rich particles compared to metal poor and solar metallicity particles. The metal rich particles lose their cluster mass significantly faster than either the metal poor or solar metallicity particles, suggesting that these cluster environments are being disrupted much more rapidly. This disruption is likely due to tidal interactions with nearby giant molecular clouds, which can distort a clusters and cause stars on the edge to be lost to the field (e.g Kruijssen, 2011; Kruijssen et al., 2012a).

The results show that as metallicity decreases, the initial fraction of mass in a cluster environment also decreases. Galaxies 4, 10, 12, 20, and 23 fully follow the proposed trend, with metal poor particles having the lowest initial fraction of mass in a cluster environment, and metal rich having the most.

If groups with less than 100 particles are removed, galaxies 6, 11, 14, and 24 do follow the full trend. This may suggest groups with small numbers of particles do not produce accurate results.

Galaxy 1 has all metal rich groups with the highest initial percentage of mass in a cluster environment. This is followed by old solar metallicity and solar type particles. Old metal poor particles have 31.2 % of their initial mass in a cluster environment, which is higher than young solar metallicity particles, which only contain 31.1 %. The other groups in this galaxy all follow the previously identified trend.

Galaxy 2 also has metal rich groups with the highest initial percentage of mass in a cluster environment. However, old metal poor particles have 34.3% of their initial mass in a cluster environment, whereas solar particles and young solar metallicity particles only have 34.1% and 31.9% respectively. This is the only anomaly in this galaxy, with the rest of the groups following the trend.

Galaxy 13 is similar to galaxy 2, with all the groups following the trend apart from old metal poor particles, which contain 31.0% of their initial mass in clusters. This is larger than the values found for young solar metallicity and solar particles, which are 30.2% and 29.1% respectively.

Galaxies 3, 5, 15, and 22 do not follow the trend. However, all of these galaxies have young metal rich particles with the highest percentage of mass in a cluster environment. In each galaxy, the remaining groups all have very similar amounts of initial mass in a cluster environment. The remaining groups in galaxy 5 have the lowest separation of 1.8%, with the largest separation of 6.7% found in galaxy 22. These separations are noticeably smaller than the previously mentioned galaxies.

As explained in Section 4.4.1, the link between metallicity and the initial fraction of mass in a cluster environment is due to a combination of physical and simulation effects, including a temperature floor.

5.4.2 Trends in Cluster Disruption Time as a Function of Metallicity for Old Particles

As mentioned in Section 4.4.2, the cuts utilised in this work will impact the results of cluster disruption. The upper age limit for both young and solar particles may mean that these groups are not disrupted by the end of the simulation, so these groups are analysed separately. The method used is the same as that outlined in Section 4.4.2, and the results are shown in Table 5.12.

TABLE 5.12: Details for the mean cluster disruption times for old star particles in galaxies. Values in italics are samples containing less than 100 particles.

Galaxy	Disruption Time (Gyrs) or Initial Mass Remaining (%)		
Number	Old Metal Rich	Old Solar Metallicity	Old Metal Poor
1	1.5	9.0	12.5
2	2.5	9.0	0.01%
3	*2.5*	6.0	10.5
4	*1.5*	9.5	0.01%
5	1.5	7.5	4.5
6	*4.5*	0.11%	0.01%
10	5.5	9.0	11.5
11	n/a	9.0	0.01%
12	2.5	9.0	12.5
13	7.5	8.5	2×10^{-4} %
14	*1.5*	8.0	13.0
15	n/a	9.0	13.0
20	n/a	8.0	0.01%
22	7.0	10.0	12.5
23	2.5	5.5	5.5
24	*1.5*	8.0	13.0

As with the halo analysis, the values in Table 5.12 show the first time step where the cluster mass has dropped to 0. As the interpolation was conducted at 0.5 Gyr intervals, each particle could have lost all its cluster mass at any time within the last 0.5 Gyrs, meaning the error on each measurement is -0.5 Gyrs.

It is again found that old metal poor particles take longer to disrupt. The exceptions to this are galaxies 5 and 23, where old solar metallicity particles take the same amount of time or longer to disrupt than the old metal poor particles. In galaxy 6, neither the old solar metallicity or old metal poor particles undergo disruption.

Overall, 13 of the galaxies follow the trend, suggesting that clusters with a large number of metal poor stars are likely to take longer to disrupt than metal rich clusters.

The mechanisms responsible for cluster disruption are outlined in Section 2.7.2.1 and the references therein, and are linked to these results in Section 5.4.7.

5.4.3 Trends in Cluster Disruption Time as a Function of Metallicity for Solar Age Particles

As solar age particles have an age limit of 4.3 - 4.7 Gyrs, and the data has been interpolated in 0.5 Gyr intervals, any groups with mass still left in a cluster environment 4.5 Gyrs after formation are considered not disrupted. It should be noted that this is not an indication that these groups will not disrupt in the future, but as the EMP simulations only evolve to z = 0, it is not possible to determine if or when these groups become fully disrupted. However, it is possible to examine the amount of mass remaining in these cluster environments, which can indicate the amount of disruption that has already occurred.

The method used to calculate the disruption time for solar age particles is the same as that outlined in Section 4.4.2, and the error on all results is -0.5 Gyrs. The results are shown in Table 5.13.

TABLE 5.13: Details for the mean cluster disruption times for young star particles in galaxies. Values in italics are samples containing less than 100 particles.

Galaxy	Disruption Time (Gyrs) or Initial Mass Remaining (%)		
Number	Solar Rich	Solar	Solar Poor
1	4.5	0.45%	n/a
2	1.5	*1.5*	*0.5*
3	2.0	3.5	n/a
4	*1.5*	0.52%	*1.0*
5	1.0	1.0	n/a
6	*1.0*	0.41%	*1.0*
10	1.0	0.08%	*4.5*
11	n/a	*4.05%*	6.09%
12	1.0	0.09%	0.50%
13	1.5	0.69%	*4.5*
14	n/a	0.35%	*1.5*
15	*0.5*	*0.06%*	2.02%
20	n/a	1.88%	2.14%
22	3.5	2.08%	4.5
23	1.0	1.5	n/a
24	*0.5*	0.49%	*4.5*

In Section 5.4.2 it was generally found that metal poor particles take the longest time to disrupt, with metal rich particles disrupting quickest. No galaxy has particles in all three groups and follows this trend.

Galaxies 3 and 23 both have metal rich particles disrupting before solar metallicity particles, but both groups lack any metal poor particles. Similarly, galaxy 1 has metal rich particles disrupting, solar metallicity particles failing to disrupt, and no metal poor particles. Therefore, these three galaxies can all be considered to follow the trend.

Galaxies 10, 13, 22, and 24 have metal rich particles disrupting before metal poor particles, but their solar groups do not disrupt, meaning these galaxies do not follow the expected trend.

Galaxies 2, 4, 5, 6, and 14 do not follow the proposed trend. In galaxy 2, both metal rich and solar metallicity particles disrupt at 1.5 Gyrs, whereas metal poor particles disrupt after 0.5 Gyrs. In galaxies 4 and 6, metal poor particles disrupt first, followed by metal rich particles, with solar metallicity particles failing to disrupt. Galaxy 5 does not have any metal poor particles, and both solar metallicity and metal rich particle disrupt after 1 Gyr. However, as the error on these measurements is -0.5 Gyrs, it is possible that solar metallicity particles disrupt after metal rich particles. In galaxy 14, solar poor particles disrupt after 1.5 Gyrs, and solar metallicity particles fail to disrupt.

Galaxies 11, 12, 15, and 20 all have solar metallicity and metal poor particles failing to disrupt. As previously mentioned, it is possible that these groups will disrupt in the future, meaning each of these galaxies could follow the trend, but it is not possible to predict this with these simulations.

As a whole, the results in Table 5.13, show a surprisingly high number of solar groups failing to disrupt. If these results were following the proposed trend, it is expected that metal poor groups would make up the majority of the groups that failed to disrupt, followed by solar metallicity groups. However, these results show that 12 out of the 16 solar groups failed to disrupt, whilst the same is only true for 4 metal poor groups. This suggests that particles with both a solar age and metallicity are more likely to retain mass in a cluster environment that particles of a similar age but different metallicities. However, the average amount of initial mass retained by solar particles is only 0.93%, compared to 2.69% for solar age metal poor particles, suggesting that solar age metal poor particles retain more of their initial mass.

To determine the disruption times of groups that are yet to disrupt, the EMP simulations would need to be evolved in to the future.

5.4.4 Trends in Cluster Disruption Time as a Function of Metallicity for Young Particles

As young particles have an age limit of less than 4.3 Gyrs, and the data has been interpolated in 0.5 Gyr intervals, any groups with mass still left in a cluster environment 4.0 Gyrs after formation are considered not disrupted. As with the solar age particles, failure to disrupt before the end of the simulation does not mean that these groups will not disrupt in the future.

The method used to calculate the disruption time is the same as that outlined in Section 4.4.2, and the error on all results is -0.5 Gyrs. The results are shown in Table 5.14.

TABLE 5.14: Details for the mean cluster disruption times for young star particles in galaxies. Values in italics are samples containing less than 100 particles.

Galaxy Number	Disruption Time (Gyrs) or Initial Mass Remaining (%)		
	Young Rich	Young Solar Metallicity	Young Poor
1	3.5	0.32%	*0.0*
2	1.5	1.5	3.5
3	2.5	3.0	n/a
4	4.0	0.15%	*3.5*
5	1.5	2.0	n/a
6	*0.03%*	0.23%	*3.0*
10	1.5	0.09%	4.0
11	*0.5*	0.02%	1.40%
12	3.0	0.18%	0.43%
13	3.5	0.04%	0.20%
14	3.0	0.05%	4.0
15	1.5	0.01%	3.23%
20	n/a	2×10^{-3}%	0.20%
22	1×10^{-3}%	0.30%	0.16%
23	1.5	1.5	n/a
24	3.0	0.13%	3.0

Galaxies 3 and 5 both have metal rich particles disrupting before solar metallicity particles. However, neither galaxy contains any metal poor particles, so they follow the proposed trend.

Galaxies 1 and 4 have metal poor particles disrupting faster than metal rich particles, and the solar metallicity particles in both galaxies fail to fully disrupt. In galaxy 6, neither the metal rich nor solar metallicity particles disrupt, and metal poor particles disrupt after 3 Gyrs. Both galaxies 10 and 14 have metal rich particles disrupting first, followed by metal poor particles, with solar metallicity particles failing to disrupt. In galaxy 24, both metal rich and metal poor take 3 Gyrs to disrupt, but solar metallicity particles do not disrupt.

The metal poor and solar metallicity particles in galaxies 2 and 23 take 1.5 Gyrs to disrupt, meaning that these galaxies do not follow the proposed trend. Similarly, galaxies 11, 12, 13, 15, and 20 all have solar metallicity and metal poor particles failing to disrupt. Whilst these groups may disrupt in the future, it is not possible to determine if or when this will occur with these simulations. No groups disrupt in galaxy 22, again meaning that it is not possible to determine if this galaxy will eventually follow the trend if it is allowed to evolve further.

There are 2 young metal rich clusters that fail to disrupt, but each of these contains less than 0.03% of their initial cluster mass, meaning they have nearly been fully disrupted. Twelve young solar metallicity groups fail to disrupt, along with six young metal poor groups. The young solar metallicity groups all have less than 0.23% of their initial cluster mass left, with the average amount remaining being 0.094%. By comparison, the undisrupted young metal poor groups have between 0.16% and 3.23% of their initial cluster mass remaining, with the average being 0.934%. This suggests that the twelve young solar metallicity groups have been disrupted to a greater extent than the six young metal poor groups, implying that if the simulations were evolved further, the young solar metallicity groups would disrupt before the young metal poor groups.

It would be helpful if future work with the EMP simulations evolved each run into the future, as this would enable the disruption times for the young particles in each galaxy to be determined.

5.4.5 Trends In Metallicity and The Amount of Time Spent in a Cluster Environment

The trends for the average expectation value for the amount of time spent in a cluster environment were calculated using the same method as that outlined in Section 4.4.5. The values from this analysis, along with the standard error on the mean, are presented in Table 5.15.

The mean expectation value generally decreases as metallicity decreases, suggesting that metal poor stars spend less time in cluster environments than their metal rich counterparts. The exception to this is old metal poor particles, which have a mean expectation value of 4.323 Myrs.

As metal poor particles generally spend less time in a cluster environment, they will have spent less time around high mass stars. As discussed in Section 1.6, these stars produce metals via nucleosynthesis, and these metals are required for the formation of

TABLE 5.15: Statistical analysis of the expectation value for the amount of time spent in a cluster environment (0-100 Myrs).

Group	Minimum (Myrs)	Maximum (Myrs)	Median (Myrs)	Mean (Myrs)	Standard Deviation
Young Metal Rich	1.14	7.71	6.92	5.70 ±0.55	2.13
Solar Age Metal Rich	1.16	8.62	7.08	5.96 ±0.61	2.19
Old Metal Rich	1.38	7.68	3.74	4.46 ±0.55	1.97
Young Solar Metallicity	1.58	7.48	3.14	3.46 ±0.42	1.69
Solar Age and Metallicity	1.45	8.42	3.85	4.14 ±0.52	2.08
Old Solar Metallicity	1.34	8.52	3.54	4.10 ±0.54	2.15
Young Metal Poor	0.0	4.42	1.96	2.06 ±0.32	1.15
Solar Age Metal Poor	0.54	3.86	2.02	2.01±0.26	0.89
Old Metal Poor	1.88	9.99	3.24	4.32 ±0.60	2.39

planets. The minimal time spent in a cluster environment therefore suggests that metal poor particles may not have the materials required to form planets.

As discussed in Section 3.2, both young and solar age particles have upper age limits which determine how long they are able to evolve for. These limits were selected as they are in line with current measurements for the age and metallicity of the Sun, meaning the results obtained in this work should not vary drastically from the results that would be obtained if this work was repeated using observational data. In addition to this, different limits were trialled for the solar group, but this made little difference to the number of particles in the solar groups. As the mean is dependent on the number of values used to calculate it, a small difference in the number of particles would not significantly change the value of the mean calculated here.

5.4.6 Trends In Metallicity and the Number of Supernova Experienced

The expectation value for the number of supernova, and the standard error on the mean, were calculated and are outlined in Table 5.16. The solar group in galaxy 11 was found to have an expectation value of over 33,000, which was significantly higher than any other values for that group. Therefore, the values for solar type particles were calculated twice: once with this outlier, and once without. The values calculated without the outlier can be found in brackets.

The results from Table 5.16 show that, when the solar results without the outlier are used, there is a decrease in the mean and median as the metallicity decreases. This suggests that metal rich particles experience more supernova in the first 100 Myrs since their formation. As it has been established that these particles also spend longer in cluster environments, where they would be nearer to more supernova, this is not unexpected.

TABLE 5.16: Statistical analysis of the expected number of supernovae experienced in the first 100 Myrs since formation for each group across all haloes.

Group	Minimum	Maximum	Median	Mean	Standard Deviation
Young Metal Rich	300	7099	3597	4266 ±560	2169
Solar Age Metal Rich	799	6506	4421	4109 ±550	1984
Old Metal Rich	963	6373	4319	3742 ±550	1982
Young Solar Metallicity	1305	5898	2560	3105 ±335	1338
Solar Age and Metallicity	1332	33604	2954	7493 ±1302	5208
	(1332)	(6773)	(2839)	(3315 ±399)	(1596)
Old Solar Metallicity	1560	6539	2822	3318 ±357	1427
Young Metal Poor	0.0	3453	1791	1857 ±254	917
Solar Age Metal Poor	85	3099	1792	1610 ±287	995
Old Metal Poor	1799	6619	2423	3100 ±344	1376

5.4.7 Conclusions

In Section 4.4.1, the initial fraction of mass in a cluster environment was shown to decrease as metallicity decreases, and this was expected to be repeated in the galaxy analysis. When the galaxies were analysed, 5 were found to fully follow the proposed trend, and another 4 would follow the trend if groups with less than 100 particles were removed. As with the results in Section 4.4.1, this suggests that metal rich particles in most of the galaxies have a higher initial fraction of mass in a cluster environment. As high metallicity environments have higher cooling rates, it is usually expected that these environments would also have lower temperatures. However, the EMP simulations have a temperature floor of 10^4K, as temperatures below this cannot be modelled. For high metallicity particles with higher cooling rates, this temperature floor means they will eventually reach a constant temperature. This means that any density increases in these particles will be accompanied by an increase in the pressure. As free-fall times are proportional to $1/\sqrt{density}$, higher pressure environments will have higher densities and shorter free-fall times (Kruijssen, 2012b). This will be accompanied by a higher star formation efficiency, meaning that a larger percentage of the mass in the particle will be rapidly converted into stars bound in cluster environments (Kruijssen, 2012b), leading to high initial fractions of mass in cluster environments for these high metallicity particles.

When habitability is considered, having a larger percentage of the initial mass in a cluster environment limits the mass available to form stars and planets in less hostile (and more habitable) environments. This means that metal rich particles are likely to host fewer stars with habitable planets than their lower metallicity counterparts.

When the time required for old clusters to disrupt was analysed, it was found that disruption time decreased as metallicity increased. This aligns with Lamers, H. J. G. L. M. et al. (2017), which finds that cluster disruption is more efficient in higher metallicity environments, such as the centre of a galaxy. This efficiency decreases as metallicity decreases, which is again seen in the results presented in Section 5.4.2. This suggests that the mass in old metal poor clusters is likely to remain in this environment for longer periods of time. Cluster environments are usually considered to be hostile to the formation of planets due to their high stellar densities (e.g Pechetti et al., 2020). However, metal poor particles were found to have the smallest mean percentage of initial mass in a cluster environment, meaning that these clusters may have lower stellar densities and may not be as hostile.

Of the 16 galaxies used in this work, 14 have less than 50% of their initial mass in a cluster environment, and 11 have less than a third of their initial mass in a cluster environment. This lack of mass would lead to a reduced stellar density and fewer supernovae in cluster environments, making them less hostile to life (Close & Pittard, 2017). This could mean that, relative to other cluster environments, more planets that form in old metal poor cluster environments survive. Whilst the resolution of these simulations does not allow for individual stars or planets to be resolved, they do allow for the fraction of mass in cluster environments to be determined. This can then be used as a proxy for the fraction of stars in dense stellar environments, providing an insight into the hostility of the cluster environments modelled in these simulations.

When cluster disruption time was analysed for solar age particles, only galaxies 1, 3, and 23 could be considered to follow the trend identified in the old particles. Of the solar groups, 12 failed to disrupt. Two of these groups contained less than 100 particles, meaning that their disruption time could have been skewed by small sample sizes. Therefore, these galaxies were not considered when the mean percentage of initial mass remaining was calculated. Only four solar age metal poor particles failed to disrupt, and none of these were impacted by small sample sizes. Upon calculating the mean percentage of initial mass remaining in a cluster environment, it was found that solar groups retained 0.702%, whereas solar age metal poor groups retained an average of 2.689%. This suggests that solar groups may be closer to being fully disrupted than solar age metal poor particles. However, confirming this would require that the EMP simulations are evolved into the future, which has not been done in this case.

For young particles, there are two young metal rich groups that fail to disrupt, but both of these have less than 0.03% of their initial mass remaining in a cluster environment, so it is likely that these groups would disrupt relatively quickly if allowed to evolve further. Twelve of the sixteen young solar metallicity group fail to disrupt, with these groups all

having less than 0.4% of their initial mass in a cluster environment remaining. Only six young metal poor groups fail to disrupt, but these groups retain significantly more of their cluster mass, with values ranging between 0.16% and 3.228%. This suggests that the young metal rich groups that have failed to disrupted would disrupt first, followed by the young solar metallicity particles, and then the young metal poor particles.

For both the solar age and young particles, it is notable that the solar metallicity groups in both cases have the fewest fully disrupted groups. This suggests that solar metallicity particles are more likely to retain mass in a cluster environment than either metal rich or metal poor particles. However, the metal poor particles that do fail to disrupt tend to retain a higher percentage of their initial cluster mass.

The large reduction of mass in a cluster environment for solar metallicity particles may reduce the stellar density, feedback, and rate of photoionisation, allowing any young planets that remain to evolve in a relatively friendly environment. As the reduction in mass was greatest for solar metallicity particles, these cluster environments will be more friendly than those in metal poor particles. This means planets forming in cluster environments in solar metallicity particles are more likely to survive when compared to those forming in cluster environments found in metal poor particles. This may suggest that stars with solar metallicities, such as the Sun, are more likely to host planets.

The analysis of the expectation values for the amount of time spent in a cluster environment suggests that metal poor stars generally spend less time in these environments. This is not unexpected given that metal poor particles have less of their initial mass in a cluster environment, but may be unexpected given the longer disruption times seen for old metal poor particles. Therefore, it must initially be pointed out that metal poor particles may also be considered as 'solar age' and 'young', so this result is not in direct contradiction to any others. It may also suggest that many stars that do find themselves in old metal poor clusters are able to move into the field relatively quickly.

Either way, as metal poor particles spend less time in cluster environments, there is a lower chance that these particles will be exposed to hostile environments for prolonged periods. This minimises the opportunities for planets around metal poor stars to be disrupted, destroyed, or undergo other processes that limit their habitability. As stars with higher metallicities spend longer in these environments, it is more likely that any planets remaining around these stars have a reduced level of habitability.

Finally, the results for the expected number of supernova in the first 100 Myrs show a decrease in mean and the metallicity decreases. Previous results have established that metal rich particles have more of their mass in a cluster environment and spend more time in a cluster environment during the first 100 Myrs of their lives. As cluster

environments host many stars that go supernova, and have higher stellar densities, it is expected that metal rich particles will experience more supernova in their early lives. As supernova are highly energetic events, being exposed to several supernova is likely to damage proto-planetary disks and planets (Close & Pittard, 2017), reducing the likelihood that metal rich stars will host any habitable planets in their later lives.

Chapter 6

Comparing Trappist and Solar Type Particles

The work thus far has focused on determining if there is anything unique about the evolutionary history of the Sun that may have enabled a habitable planet to form around it. This focus was dictated by the fact that the only known habitable planet is Earth, making the Sun the only star to host life. However, the NASA Exoplanet Archive currently lists 36 other stars that are known to host planetary systems with at least 5 confirmed planets, proving that multi-planet systems are not unique to the Sun. To determine if the evolutionary history of another star with a multi-planet system is similar to that of a solar type star, the galaxy analysis was repeated to compare particles that would represent these two stars.

6.1 Selecting A Comparison Star

As several of the solar groups have been identified as containing less than 100 particles, it was decided to eliminate any comparison particles that were classified as solar type. This limited the possibility of the new comparison group containing less than 100 particles, which in turn would limit the statistical robustness of the results, and also allowed for solar type particles to be compared to the new comparison particles. By comparing these two groups, it would be possible to identify similarities in their evolutionary histories which may have enabled the stars represented by these particles to host planets.

Apart from this, there were no inherent limits on the new target particle, so a list of additional criteria were created. This work has focused on identifying the evolutionary histories of solar type particles and how this may impact the formation of planets.

Therefore, comparing the host star of the solar system to the host star of an analogous planetary system may highlight evolutionary traits unique to these systems. Therefore, the first condition that was imposed was that the star being represented by these particles must host between 7 and 9 confirmed planets. This range was chosen as our solar system is currently home to 8 confirmed planets.

However, this method is not without potential flaws. A mission with a lifetime similar to that of Kepler would likely only detect around 3 planets orbiting a star, with the outer planets that have longer orbits likely being missed. This may mean that other, more suitable planetary systems have been discounted due to these planets being missed. However, as it is not possible to confirm if there are additional planets around a star without further observations, this method is the only way to guarantee that a star hosts between 7 and 9 planets. As more observational data is gathered and more multi- planet systems are identified, it would be beneficial for this analysis to be run again using new comparison stars.

The next criteria was that the age and metallicity ranges for the host star must fall within one of our previously defined groups (except the solar group). This would allow the comparison group to be compared to all the other groups within the galaxy, mimicking the analysis completed in previous chapters.

To determine which planetary systems met this criteria, the NASA Exoplanet Archive, which contains data for over 5000 planets, was used (Akeson et al., 2013). After limiting the results to only include stars which host between 7 and 9 planets, the only remaining stars were KOI-351 and TRAPPIST-1.

6.1.1 KOI-351

KOI-351 has 8 confirmed planets (Shallue & Vanderburg, 2018), an age of 0.53 ± 0.88 Gyrs (Burke et al., 2014), and an [M/H] metallicity of 0.12 ± 0.18 dex (Cabrera et al., 2014). This means that KOI-351 has the same number of planets as the solar system, making it a good analogue. However, the metallicity of this star is measured in [M/H] rather than [Fe/H], and whilst these values would be similar, they are not identical. When the errors on the metallicity measurement are taken into account, the metallicity range is -0.3 - 0.06 dex, meaning it spans both the young metal poor and young solar metallicity groups, so it cannot be used in this work.

TABLE 6.1: Details of the metallicity and age cuts used to classify TRAPPIST type particles.

Category	Cut Applied
Age	$5.4 \leq$ Age ≤ 9.8 Gyrs
Metallicity	$-0.04 \leq$ [Fe/H] ≤ 0.12

6.1.2 TRAPPIST-1

TRAPPIST-1 hosts 7 exoplanets (Burgasser & Mamajek, 2017), has an age of 7.6 ± 2.2 Gyrs (Burgasser & Mamajek, 2017), and an [Fe/H] metallicity of 0.04 ± 0.08 dex (Gillon et al., 2017). This again makes the TRAPPIST-1 system a reasonable analogue for the solar system. When the full errors are accounted for in both the age and the metallicity measurements, the host star falls into the old solar metallicity group. This means that TRAPPIST type stars will be a sub group of the old solar metallicity group that has bee previously identified.

In addition to the properties of the host star, the exoplanets in the TRAPPIST-1 system make this star interesting. In particular, TRAPPIST-1e is considered to be a good candidate for a habitable planet. It is rocky, orbits in the habitable zone (de Wit et al., 2018), has a mass $0.69 M_{\oplus}$, and a radius of $0.92 R_{\oplus}$ (Agol et al., 2021). This makes the planet reasonably similar to Earth, and, when combined with the properties of the host star, make TRAPPIST-1 a good system for use in this work.

The age and metallicity limits used to classify a TRAPPIST type particle are outlined in Table 6.1.

6.2 Comparisons within Galaxies

For the purpose of this analysis, TRAPPIST type particles were only compared to solar type particles. This was initially done within each individual galaxy.

6.2.1 Statistical Analysis of the Galactocentric Radius within Galaxies

The galactocentric radius at which a particle sits correlates strongly with the metallicity of the particle. If TRAPPIST type particles were found to occupy similar regions in their host galaxies to solar type particles, this may suggest that these areas are better suited for planet formation, and may therefore be more likely to host life. Both Kolmogorov-Smirnov and Anderson-Darling tests were carried out, and the same 2σ limit of -1.3 was applied. The results found to be above this cut off are outlined in Table 6.2.

TABLE 6.2: Statistically significant results of Kolmogorov-Smirnov and Anderson-Darling tests comparing the galactocentric radii of TRAPPIST and Solar type particles

Galaxy	$\log_{10}$ KS p-value	$\log_{10}$ AD p-value
11	*-0.001*	*-0.602*
15	*-0.271*	*-0.602*
20		-1.159

The Kolmogorov-Smirnov and Anderson-Darling tests each provide 16 results, but only 2 Kolmogorov-Smirnov and 1 Anderson-Darling results were above the 2σ cut off. As galaxy 20 only produces a single result rather than a pair, this is not considered when determining if there are any trends. In addition to this, both the solar and TRAPPIST groups in galaxies 11 and 15 are below 100 particles, meaning results from these groups are statistically limited.

Combined, these results suggests that solar and TRAPPIST type particles do not have similar galactocentric radius distributions. However, this result was not unexpected, as previous analysis had shown that old solar metallicity particles did not have similar radial distributions to solar type particles.

6.2.2 Statistical Analysis of the Mass-Weighted Time in Galaxies

The mass-weighted time was also compared between solar and TRAPPIST type stars using Kolmogorov-Smirnov and Anderson-Darling tests. The results found to be above the 2σ cut off are listed in Table 6.3.

TABLE 6.3: Statistically significant results of Kolmogorov-Smirnov and Anderson-Darling tests comparing the mass-weighted time in cluster environments of TRAPPIST and Solar type particles.

Galaxy	$\log_{10}$ KS p-value	$\log_{10}$ AD p-value
2	*-0.598*	*-0.624*
15	*-0.536*	*-0.602*
20	-0.400	-0.602

As with the galactocentric radius analysis, there were sixteen results from each of the Kolmogorov-Smirnov and Anderson-Darling tests. However, only 3 values from each test were above the cut off, and only one from each test was not limited by a small sample size. Both results that are not limited by a small sample size are found in galaxy 20. This suggests that solar and TRAPPIST type particles in this galaxy share similar mass weighted time distributions. However, as this is not seen in any other galaxy, it is likely that this is due to the initial conditions of this particular galaxy, rather than due to a fundamental part of the evolutionary histories of these particles. Overall, these

results suggest that the TRAPPIST and solar distributions of MWT are not similar in most galaxies.

6.2.3 Statistical Analysis of the Number of Supernovae Experienced in the First 100 Myrs Since Formation

The expected number of supernovae in the first 100 Myrs was also analysed. Results above the 2σ cut off are listed in Table 6.4.

TABLE 6.4: Statistically significant results of Kolmogorov-Smirnov and Anderson-Darling tests comparing the number of supernovae experienced in the first 100 Myrs since formation of TRAPPIST and Solar type particles.

Galaxy	$\log_{10}$ KS p-value	$\log_{10}$ AD p-value
2	*-0.352*	*-0.602*
11	*-0.550*	*-0.616*
15	*-0.414*	*-0.602*
20	-0.014	-0.602

Only 4 results above the cut off were obtained for each of the 2 tests. However, 3 results from each test are limited by small sample sizes. Again, the results from galaxy 20 are the only values not limited by a small sample size. This suggests that these particles experience similar numbers of supernovae in their early lives. However, as this result is only seen in four of the galaxies, this may be down to the initial conditions of these particular galaxies, rather than a fundamental property of both TRAPPIST and solar type particles.

6.2.4 Statistical Analysis of the Amount of Time Spent in a Cluster Environment in the First 100 Myrs Since Birth

As with our previous analysis, Kolmogorov-Smirnov and Anderson- Darling tests were used to compare the amount of time spent in a cluster environment in the first 100 Myrs after formation. The results found to be above the 2σ cut off are listed in Table 6.5.

TABLE 6.5: Statistically significant results of Kolmogorov-Smirnov and Anderson-Darling tests comparing the amount of time spent in cluster environments of TRAPPIST and Solar type particles.

Galaxy	$\log_{10}$ KS p-value	$\log_{10}$ AD p-value
2	*-1.282*	*-0.917*
13	-0.433	
15	*-0.565*	*-0.602*

As galaxy 13 only has a Kolmogorov-Smirnov test above the 2σ cut off, this result is not considered when identifying potential trends. The results from galaxies 2 and 15 are all

limited by small sample sizes, meaning any conclusions drawn from these results may be limited. As only two galaxies produce a pair of results, it is unlikely that TRAPPIST and solar type particles spend similar amounts of time in a cluster environment in their early lives.

6.2.5 Conclusions

The analysis carried out on individual galaxies does not suggest any strong correlation between the solar group and TRAPPIST type particles.

However, galaxy 20 produced the only results that were not limited for both mass-weighted time and the number of supernova experienced in the first 100 Myrs since formation. This suggests that the evolutionary histories of TRAPPIST and solar type particles are more similar in this galaxy compared to the others. This galaxy also produced a positive Kolmogorov-Smirnov result when the galactocentric radius distributions were compared. However, there was not a complementary Anderson-Darling result, and the Kolmogorov- Smirnov result was close to the 2σ cut off. This suggests that whilst there are some similarities between the two groups in this galaxy, elements of their evolutionary histories differ.

Overall, these results suggest that TRAPPIST and solar type particles have different evolutionary histories in these simulations. This could suggest that multi-planet systems with solar metallicity can form in a range of different environments, and do not depend too heavily on the evolutionary history of their host star. However, as the resolution of these simulations do not allow planets to be resolved. this cannot be confirmed in this work.

It is also possible that TRAPPIST and solar type stars are unusual within their own age-metallicity groups. As neither the Sun or TRAPPIST has been explicitly included in these simulations, the results are based on particles with similar properties. However, this does not rule out the possibility that the Sun and TRAPPIST have different analogues to the particles used here, which could explain the lack of similarity in the evolutionary histories between the two groups.

6.3 Establishing Links Between Metallicity and Evolutionary History

As with both the halo and galaxy analysis, the initial fraction of mass in a cluster environment, the cluster disruption time, the amount of time spent in a cluster environment,

and the number of supernovae experienced in the first 100 Myrs were all analysed. However, as the aim of this section is to determine if there are any similarities between two types of planet-hosting stars, only solar type, TRAPPIST type, and old solar metallicity particles will be compared. The first reason for limiting the comparison is to determine if there are any similarities between the solar and TRAPPIST type particles, as this may highlight something that has potentially helped the formation of planets and increased the possibility of life being present. The second reason is to enable any differences between TRAPPIST type particles and old solar metallicity particles to be highlighted. As TRAPPIST type particles are a subgroup of old solar metallicity particles, any obvious differences may highlight a difference in the evolutionary histories of TRAPPIST type particles compared to other old solar metallicity particles. These differences could suggest a possible characteristic that has enabled TRAPPIST to form planets.

6.3.1 Initial Fraction of Mass in a Cluster Environment

The initial fraction of mass in a cluster environment was calculated using the method described in Section 3.5. The mean values are detailed in Table 6.6, and the errors given are the standard error on the mean.

TABLE 6.6: The mean initial percentage of mass in a cluster environment for solar, TRAPPIST, and old solar metallicity in all galaxies.

Galaxy Number	Mean % of Initial Mass in Cluster Environments		
	Old Solar Metallicty	Solar	TRAPPIST
1	33.6 ±0.1	32.9 ±0.3	27.4 ±0.2
2	41.2 ±0.1	*34.1*±3.2	37.6 ±0.1
3	46.5 ±0.2	45.9 ±0.3	29.1 ±0.4
4	22.9 ±0.1	24.8 ±0.3	23.5 ±0.2
5	57.0 ±0.2	58.3 ±0.5	43.9 ±0.3
6	19.9 ±0.2	18.7 ±0.7	19.5 ±0.5
10	48.3 ±0.1	47.6 ±0.3	33.0 ±0.1
11	18.5 ±0.2	*23.8* ±3.4	*18.9* ±1.5
12	36.3 ±0.1	40.9 ±0.3	28.5 ±0.2
13	35.5 ±0.1	29.1 ±0.6	24.5 ±0.1
14	21.4 ±0.1	20.9 ±0.3	21.1 ±0.3
15	20.5 ±0.6	*20.1* ±2.2	*15.6* ±2.7
20	18.2 ±0.2	19.7 ±0.7	20.4 ±1.1
22	16.4 ±0.1	15.9 ±0.2	16.5 ±0.1
23	52.9 ±0.1	53.8 ±0.8	36.2 ±0.2
24	24.3 ±0.1	25.9 ±0.2	22.3 ±0.4

From Table 6.6, there 11 pairs of results that are consistent with each other when errors are considered. In galaxies 6, 11, 14, and 22 the old solar metallicity and the TRAPPIST groups have similar mean percentages of initial mass in cluster environments. As the

TRAPPIST group is a subgroup of the old solar metallicity group, it is expected that these particles will have had similar early evolutionary histories, so this result is not unexpected.

In galaxies 6, 11, 14, 15, and 20, solar and TRAPPIST particles show similarities in the mean initial percentage of mass in a cluster environment. This suggests that in these galaxies, TRAPPIST and solar type particles had similar distributions of mass between cluster environments and the field. As particles are representative of stellar populations, this may suggest that both the Sun and TRAPPIST were formed in similar early evolutionary environments. This may have provided the materials needed to form planets without causing their destruction. However, due to the resolution limits, this would need to be confirmed in future work. It should also be noted that the solar and TRAPPIST groups in galaxies 11 and 15 all contain less than 100 particles, which may limit the accuracy of these results.

Additionally, galaxies 15 and 23 show similar mean initial percentages of mass in cluster environments for their solar and old solar metallicity groups. Again, it should be noted that the solar group from galaxy 15 contains less than 100 particles, potentially reducing the accuracy of this result. This suggests that solar and old solar metallicity particles in these galaxies could have experienced similar early evolutionary histories. As previous results in this work have shown similarities between the solar groups and groups with either a solar age or metallicity, these results are not unusual. It is likely that the similarity in the metallicities of these particles is representative of similar densities in the early stages of cluster formation.

6.3.2 Comparison of Cluster Disruption Time for Solar, TRAPPIST, and Old Solar Metallicity Particles

The disruption time for solar and old solar metallicity particles had already been calculated for the analysis in Sections 5.4.3 and 5.4.2 respectively. The disruption time for TRAPPIST type particles was calculated using the method outlined in Section 4.4.2. As both TRAPPIST and old solar metallicity particles are classified as old particles, the only age limit imposed on them is from the age of the simulation. However, solar particles have an upper age limit of 4.7 Gyrs, meaning these particles may not have enough time to fully disrupt before the simulations ends. Therefore, if particles have fully disrupted by the end of the simulation, Table 6.7 lists the time at which they are first recorded as containing no mass in a cluster environment. If the particles fail to disrupt, Table 6.7 lists the percentage of their initial cluster mass that remains in a

cluster environment at the end of the simulations. For the disruption times, the errors are all -0.5 Gyrs.

TABLE 6.7: Details for the cluster disruption times or remaining percentage of initial mass in cluster environments for old solar, solar, and TRAPPIST type star particles in galaxies. Values in italics are samples containing less than 100 particles.

Galaxy Number	Disruption Time (Gyrs) or Initial Mass Remaining (%)		
	Solar	Old Solar Metallicity	TRAPPIST
1	0.448%	9.0	6.5
2	*1.5*	9.0	7.0
3	3.5	6.0	3.0
4	0.517%	9.5	7.0
5	1.0	7.5	3.0
6	0.410%	0.106%	6.5
10	0.082%	9.0	6.5
11	*4.050%*	9.0	*6.5*
12	0.086%	9.0	7.0
13	0.685%	8.5	7.0
14	0.347%	8.0	6.5
15	*0.062%*	9.0	*1.0*
20	1.878%	8.0	5.5
22	2.075%	10.0	7.0
23	1.0	5.5	1.5
24	0.488%	8.0	5.0

As galaxies 11 and 15 have less than 100 TRAPPIST type particles, the results from these groups may have been impacted by the small sample sizes, so they are not considered when calculating any averages or establishing trends.

From Table 6.7 it can be seen that TRAPPIST type particles always disrupt faster than when the old solar metallicity particles are considered as a whole. The average disruption time for TRAPPIST type particles is 5.6 Gyrs, whereas old solar metallicity particles take an average of 8.3 Gyrs to disrupt. This suggests that cluster disruption mechanisms are stronger in TRAPPIST type particles, which will enable stars in TRAPPIST type particles to move into the field faster than those in old solar metallicity particles.

Due to the age limit imposed on solar type particles, it is difficult to draw a direct comparison between this group and the two others. However, on the three occasions that a solar group with more than 100 particles has disrupted, the difference in the disruption time between the solar and TRAPPIST groups are between 0.5 and 2.0 Gyrs. In all three cases, the disruption time of the TRAPPIST particles has been closer to that of the solar particles compared to the old solar metallicity particles. This may suggest that if the simulations were able to evolve into the future, TRAPPIST and solar type particles would have more similar disruption times than TRAPPIST and old solar

metallicity particles. This would suggest that these types of particles retain cluster mass for a similar amount of time. This would need to be examined in future studies.

6.3.3 Trends in Metallicity and the Time Spent in a Cluster Environment

As previously discussed, cluster environments can be hostile to the formation of planets, meaning that stars spending long periods of time in a cluster environment may not be able to host planets. To determine if planet-hosting stars have similar evolutionary histories, the expectation values for the amount of time in a cluster environment was compared for solar, TRAPPIST, and old solar metallicity particles across all galaxies. These values are shown in Table 6.8. Although the comparison being made is between solar and TRAPPIST type particles, old solar metallicity particles have been included as TRAPPIST particles also fall into this group.

TABLE 6.8: Expectation values for the amount of time spent in a cluster environment for the first 100 Myrs since formation for solar, old solar metallicity, and TRAPPIST type particles in each galaxy.

Galaxy	Solar (Myrs)	Old Solar Metallicity (Myrs)	TRAPPIST (Myrs)
1	4.330 ±0.075	4.072 ±0.033	3.507 ±0.033
2	*4.380* ±0.516	6.153 ±0.032	6.275 ±0.040
3	6.855 ±0.068	5.762 ±0.070	2.888 ±0.086
4	3.579 ±0.090	2.502 ±0.017	3.313 ±0.032
5	6.996 ±0.142	5.679 ±0.052	4.044 ±0.042
6	1.910 ±0.131	1.734 ±0.034	2.025 ±0.074
10	6.167 ±0.060	7.659 ±0.030	5.256 ±0.028
11	*1.724* ±0.387	2.417 ±0.051	*2.545* ±0.211
12	5.636 ±0.063	4.954 ±0.023	4.200 ±0.034
13	3.964 ±0.114	5.133 ±0.021	4.015 ±0.021
14	3.019 ±0.060	2.630 ±0.028	3.307±0.063
15	*2.237* ±0.286	2.162 ±0.140	*2.918* ±0.515
20	1.860 ±0.092	1.882 ±0.043	2.500 ±0.177
22	1.445 ±0.026	1.339 ±0.010	2.123±0.015
23	8.418 ±0.157	8.518 ±0.047	6.528 ±0.062
24	3.729 ±0.050	3.005 ±0.031	3.571 ±0.091

When comparing the TRAPPIST expectation values to those of the solar and old solar metallicity groups in each galaxy, 10 galaxies (4, 6, 10, 13, 14, 15, 20, 22, 23, and 24) showed a greater level of similarity between the solar and TRAPPIST expectation values. This suggests that in these galaxies, TRAPPIST and solar type particles spend broadly similar amounts of time in cluster environments in their early evolutionary stages.

To determine the significance of these similarities, the errors on each expectation value needed to be accounted for. If the inclusion of the errors showed an overlap between the expectation values from two groups, it is likely that these groups spent very similar amounts of time in cluster environments. If there was no overlap, it suggests that whilst the TRAPPIST group may be more similar to the solar group, the similarity is not significant. Out of the 10 galaxies that showed a greater level of similarity between TRAPPIST and solar groups, only 3 galaxies (6, 13, and 15) had an overlap between the errors. However, the expectation values in these 3 galaxies are not similar to each other, meaning this result does not indicate an 'ideal' amount of time to spend in a cluster environment. It is also notable that galaxy 15 has fewer than 100 solar and TRAPPIST particles, meaning that the errors on both measurements are very large. As this overlap is being used to determine if previously identified similarities are significant, this result is likely to have been impacted by the large errors.

Similarly, of the 6 galaxies that show a greater level of similarity between TRAPPIST and old solar metallicity groups, only galaxy 11 has an overlap between the errors on the measurements. This result may be indicative of the TRAPPIST and old solar metallicity particles in galaxy 11 spending similar amounts of time in a cluster environment. However, as this galaxy has less than 100 particles for the TRAPPIST group, this result may have been impacted by the small sample size.

The overall lack of overlap between TRAPPIST groups and solar groups suggest that there is often minimal similarity between the amount of time these groups spend in a cluster environment.

6.3.4 Trends in Metallicity and the Expected Number of Supernovae

The number of supernovae experienced in the first 100 Myrs since formation is calculated using the method outlined in 3.7, and the results are outlined in Table 6.9. The errors are all calculated by dividing the standard deviation my the square root of the number of particles in the group, then rounding this to the nearest whole number of supernovae. The solar group in galaxy 11 had a mean approximately 10 times larger than the other galaxies. This was due to several of the 44 particles experiencing tens of thousand of supernovae events in the first 100 Myrs. Due to the lack of solar type particles, these outliers have skewed the results given, so galaxy 11 is not considered when drawing conclusions about these results.

When the expectation values of the TRAPPIST group in each galaxy are compared to those of the solar and old solar metallicity groups, galaxies 1, 2, 3, 4, 6, 13, 15, 20, and 22 show a greater level of consistency between the TRAPPIST and solar groups.

TABLE 6.9: Expectation values for the number of supernovae experienced in the first 100 Myrs since formation for solar, old solar metallicity, and TRAPPIST type particles in each galaxy.

Galaxy	Solar	Old Solar	TRAPPIST
1	2839 ±93	3009 ±72	2144 ±86
2	*3123* ±602	4204 ±76	3332 ±59
3	4552 ±79	4986 ±113	3051 ±223
4	2134 ±255	2359 ±231	1819 ±91
5	6773 ±162	6539 ±96	5049 ±201
6	1332 ±175	2237 ±226	1619 ±481
10	5212 ±83	4614 ±30	3445 ±125
11	*33604* ±21683	2124 ±176	*1759* ±763
12	4464 ±180	3684 ±68	2881 ±175
13	3068 ±386	3673 ±124	2763 ±283
14	2329 ±461	2133 ±116	1507 ±175
15	*1487* ±397	1560 ±220	*692* ±311
20	1658 ±322	1578 ±132	1648 ±636
22	2572 ±692	2636 ±196	2416±358
23	5722 ±185	5369 ±152	3105 ±107
24	2462 ±119	2377 ±92	1304 ±93

Galaxies 5, 10, 12, 14, 23, and 24 show a greater level of consistency between the TRAPPIST and old solar metallicity groups. This initially suggests that the TRAPPIST and solar particles show more similarity in the number of supernovae experienced in their early lives compared to TRAPPIST and old solar metallicity particles. However, when the standard error on the mean is taken into consideration, the picture becomes more complicated.

After accounting for the errors, galaxies 2 and 13 show an overlap between the solar and TRAPPIST groups, suggesting that these groups experience a similar number of supernovae in their early lives. However, the solar group in galaxy 2 contains less than 100 particles, which has caused a large error and may have skewed the expectation value.

In galaxies 5, 14, 15, and 24, there are better overlaps between the solar and old solar metallicty groups, suggesting that in these galaxies, old solar and solar groups experience similar numbers of supernovae in their early lives.

In galaxies 20 and 22, all 3 groups overlap with each other, indicating that there is no significant difference in the number of supernovae experienced by any group. However, if only the calculated expectation values are considered, galaxy 20 shows more similarity between the solar and TRAPPIST groups, whereas galaxy 22 shows more similarity between the solar and old solar metallicity group. In both cases, the additional similarity is limited, and it is likely that all three groups in each galaxy experience a similar number of supernovae in their early lives.

Galaxy 4 also shows multiple overlaps, with the solar group overlapping with both the old solar metallicity and TRAPPIST groups. There is a greater level of overlap between the old solar metallicity and the solar groups, suggesting a marginally higher level of similarity between them. However, the expectation values for all three groups are consistent with each other, so the difference in the number of supernovae experienced is likely to be minimal.

In galaxy 6, there is overlap between TRAPPIST particles and both solar, and old solar metallicity particles, but no overlap between the solar and old solar metallicity groups. Whilst this suggests that the TRAPPIST particles experience a similar number of supernovae during their early evolution as both the solar and old solar metallicity groups, the expectation values are more similar between the solar and TRAPPIST particles. This suggests that TRAPPIST and solar particles have more similar early lives.

For galaxies 1, 3, 10, and 23, there are no overlaps between any groups, even when the errors are considered. However, all four galaxies show a greater level of similarity between the solar and old solar metallicity groups. This suggests that solar and old solar metallicity particles may have more similar distributions in the number of supernovae, but this is not significant.

For galaxy 12, there is no overlap between any of the 3 groups, and none of the values show any significant similarity. This suggests that there is no real similarity between the number of supernovae in the early lives of the three groups in considered from this galaxy.

The initial comparison of the expectation values suggests that TRAPPIST and solar type particles show more similarity in the number of supernovae experienced in the first 100 Myrs. However, when overlaps between the errors on the expectation values are considered, only 6 galaxies have overlaps between the solar and TRAPPIST groups, and only 4 of these represent the greatest level of similarity between two groups in the given galaxy. Similarly, there are 3 galaxies in which the TRAPPIST and old solar metallicity groups have overlaps, but none of these represent the greatest level of similarity between two groups in the given galaxy.

Overall, these results suggest that the number of supernovae experienced by TRAPPIST particles is most similar to the number of supernovae experienced by solar particles. As supernovae are capable of producing the materials needed for planet formation, and both the Sun and TRAPPIST-1 host planets, it is possible that the similarity in the expectation values indicates an 'ideal' number of supernovae. However, whilst the values are similar within some of the galaxies, they are not similar between different galaxies,

suggesting TRAPPIST particles in different galaxies do not experience similar numbers of supernovae.

6.3.5 Conclusions

The results from the comparisons of the expectation values show several areas of similarity between solar and TRAPPIST particles.

Despite being a subgroup of the old solar metallicity group, the TRAPPIST particles showed a greater level of similarity to solar particles when the mean percentage of initial mass in a cluster environment was considered. This suggests that TRAPPIST and solar type stars may begin their lives in clusters of similar masses, which in turn may have provided similar early environmental conditions that impact the possibility of habitable planets being formed.

TRAPPIST type particles also show a notable difference in disruption time compared to old solar metallicity particles. TRAPPIST particles disrupt an average of 2.7 Gyrs faster than the old solar metallicity group as a whole, suggesting that TRAPPIST type stars leave their host clusters earlier than old solar metallicity stars.

Due to the upper age limit used for solar type particles, there were limited opportunities to draw comparisons between the solar and TRAPPIST groups. However, on the occasions it was possible, the TRAPPIST groups were found to have disruption times closer to that of the corresponding solar group, suggesting that cluster disruption for both particle types occurs on similar scales. As it has already been established that TRAPPIST and solar type particles have similar percentages of their initial mass in cluster environments, this is not unexpected.

The amount of time spent in a cluster environment in the early lives of both TRAPPIST and solar type particles shows some similarity, but this is often insignificant, suggesting that these two groups rarely spend very similar amounts of time in these environments. The results highlighted 3 galaxies where there was a significant overlap between the two groups, but as this trend is not repeated across all galaxies, it is likely to be the result of different initial conditions.

When considering the number of supernovae experienced in the first 100 Myrs since formation, solar type particles initially show more similarity to TRAPPIST particles, with 9 galaxies having TRAPPIST expectation values closer to solar group expectation values than those of old solar metallicity groups. However, when the errors are considered, only 6 galaxies show overlap between TRAPPIST and solar groups, and of these 6 only 2 of the overlaps indicate the strongest similarity between groups within their

host galaxy. Old solar metallicity particles have overlaps with TRAPPIST groups in 4 galaxies, but none of these overlaps represent the highest level of similarity within the host galaxy.

These results suggest that whilst TRAPPIST groups show more similarity to solar groups, this is not seen across all galaxies, and is therefore likely to be the result of initial conditions. As the levels of similarity between TRAPPIST and solar groups are minor, these results suggest that the number of supernovae experienced in the first 100 Myrs is unlikely to be the most important factor when establishing how likely a particle is to form planets.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

7.1.1 Conclusions from the Halo Analysis

The work carried out in Section 4.2 on individual haloes demonstrates that particles with both a solar age and metallicity have unique distributions for their halo-centric radii, mass-weighted time, expected number of supernovae, and amount of time spent in a cluster environment. This suggests that solar type particles, which are used as proxies for the Sun, have different evolutionary histories when compared to particles from the same halo with different ages and metallicities.

However, when the solar groups were compared to each other using statistical tests (see Section 4.3), no strong similarities in the results were found. This suggests that solar particles from different haloes have different evolutionary histories, and that not all solar type particles experience the same environments throughout their lives. This is likely to be due to the different initial conditions imposed for each halo, and suggests that the evolutionary history of a Sun type star is highly environment dependent.

From the halo analysis in Section 4.4.5, it was determined that metal poor particles spend longer in cluster environments. Cluster environments are usually considered to be hostile to planet formation, as these environments usually have high stellar densities (e.g Pechetti et al., 2020), increasing the threat from photoionisation and supernovae. However, the analysis in Sections 4.4.1 and 4.4.6 shows that metal poor particles experienced fewer supernovae, and contained less of their initial mass in a cluster environment. This would lead to lower stellar densities within metal poor cluster environments, meaning they may not be as hostile to planet formation as is usually assumed.

When the cluster disruption times of the different old groups were examined (see Section 4.4.2), the particles showed an increase in disruption time as metallicity decreased. This suggests that old particles with higher metallicities lose all their clusters much faster than their lower metallicity counterparts, forcing all stars to continue their evolution in the field environment. As field environments tend to be less hostile than their cluster counterparts, moving into the field quickly may enable old, metal rich particles to form and retain planets more easily than lower metallicity particles of the same age.

When solar age groups were examined in Section 4.4.3, more groups failed to disrupt all their clusters due to the limit placed on their maximum age. However, it was expected that these groups would still show signs of following the trend seen in the old particles, meaning that solar age metal poor groups were expected to account for most of the groups that failed to disrupt. Upon examination, two solar age metal rich, seven solar age metal poor, and thirteen solar age and metallicity groups failed to disrupt. This suggests that solar particles retain mass in a cluster environment for longer periods than particles of the same age and different metallicities. However, as 12 of the solar groups that had failed to disrupt retained less than 4% of their initial cluster mass, it is likely that these environments would not have been particularly hostile to planet formation and evolution. This may suggest that if planets were able to survive the more hostile stages of cluster evolution, they may now be able to thrive in these environments.

The analysis of young particles in Section 4.4.4 was also expected to show signs of following the trend seen in the old particles. However, 5 metal rich, 8 metal poor, and 13 solar metallicity groups failed to disrupt, again suggesting that solar metallicity groups retain their mass in cluster environments for longer than particles of the same age but different metallicities. Young solar metallicity particles only retained an average of 0.19% of their initial cluster mass, compared to 1.05% for young metal poor particles. This again suggests a low stellar density in these environments, potentially making them less hostile to planet formation and survival. These averages also suggest that if the simulations were evolved further, young solar metallicity particles would disrupt before the young metal poor particles, replicating the pattern seen in the old particles.

7.1.2 Conclusions from the Galaxy Analysis

The statistical tests conducted on a galaxy level showed no significant similarities between solar particles and other particles within their host galaxies (see Section 5.2). This was also the case when solar particles from different galaxies were compared (see Section 5.3). This suggests that the evolutionary histories of solar type particles are not

uniform across different galaxies, and are instead dependent on environmental factors governed by the initial conditions used for each galaxy.

Section 5.4.6, shows that the expectation values for the number of supernovae experienced in the first 100 Myrs for solar particles did show some similarities between certain galaxies. However, this was not repeated across all galaxies, suggesting that this is due to the initial conditions rather than a property of the particles themselves.

Section 5.4.1 demonstrated that the initial fraction of mass that was found within a cluster environment decreased with metallicity on both a halo and galaxy scale. This suggests that metal rich particles have the highest initial fraction of mass in cluster environments. As cluster environments usually have higher stellar densities (e.g Pechetti et al., 2020), planets forming in these environments are likely to be exposed to high levels of photoionisation, which may disrupt or destroy them.

As previously found in the halo analysis, disruption time for old clusters (Section 5.4.2) in galaxies was found to increase as metallicity decreased. For particles with a solar age (Section 5.4.3), all the metal rich groups disrupted, but 12 of the solar groups and 4 solar age metal poor groups failed to disrupt. This again seems to contradict the trend first identified in the old particles from the halo analysis, which would suggest that more solar age metal poor groups would fail to disrupt. However, upon closer examination, the remaining solar groups only retained 0.70% of their initial cluster mass, compared to 2.69% for solar age metal poor groups. This suggests that the disruption process is likely to finish earlier in the solar groups, but confirming this would require the EMP simulations to be evolved into the future.

When the young groups were analysed within their respective galaxies in Section 5.4.4, 2 metal rich, 12 solar metallicity, and 6 metal poor groups fail to disrupt. The metal rich groups both contain less than 0.03% of their initial cluster mass, with the solar metallicity groups retaining less than 0.4%. The metal poor groups have retained between 0.16% and 3.23% of their initial mass.

For both the solar age and young particles, the solar metallicity groups have the fewest fully disrupted groups. This suggests that solar metallicity particles are more likely to retain mass in a cluster environment. However, the metal poor groups that do not disrupt retain a higher percentage of their initial cluster mass. In both cases, the reduction of mass in these cluster environments may reduce the stellar density, making these environments less hostile to any planets that remain.

The expectation values for the amount of time spent in a cluster environment is calculated in Section 5.4.5. This shows that metal poor particles generally spend less time in these environments. This suggests that stars forming in old metal poor clusters are able

to move into the field relatively quickly, reducing their exposure to destructive events. This may allow protoplanetary disks and young planets to survive around these stars. By contrast, higher metallicity stars spend longer in these environments, meaning that planets hosted by theses stars are at a greater risk of being destroyed or damaged by supernovae and photoionisation.

The mean expected number of supernovae in the first 100 Myrs in galaxies (Section 5.4.6) show the same decrease with metallicity as seen in the halo analysis. As metal rich clusters have been shown to have higher initial fractions of masses in cluster environments, it is likely that these supernovae are occurring in clusters. This means that any protoplanetary disks in metal rich cluster environments are subjected to more supernovae, preventing the formation of habitable planets (Close & Pittard, 2017).

Overall, the halo and galaxy analysis gave very similar results. This is not unexpected, as all particles contained in the galaxies were also contained in their associated haloes.

7.1.3 Conclusions from the TRAPPIST Analysis

The comparison of TRAPPIST and solar type particles showed up several similarities, particularly when the properties considered focused on clusters.

Both groups have similar mean percentages of their initial mass in cluster environments (Section 6.3.1), and both spend similar amounts of time in a cluster environment (Section 6.3.3). Combined, these results suggest that the Sun and TRAPPIST had similar early evolutionary histories, which may have determined how many planets they were able to form, and the chemical make up of these planets.

In addition to this, TRAPPIST type particles shared very few similarities to old solar metallicity particles, despite being a subgroup of this particle type. In Section 6.3.2, it was found that cluster environments in TRAPPIST type particles disrupt significantly quicker than those in old solar metallicity particles, although the upper age limit placed on solar particles means it is not possible to conclusively determine if they are more similar to TRAPPIST particles.

Since both the Sun and TRAPPIST are known to host multi-planet systems, it is possible that the similarities between these groups are the result of both stars residing in conditions that encourage planet formation and survival. However, as TRAPPIST is currently the only star that has been used as a comparison, this cannot be said conclusively.

7.1.4 Constraining n_e in Drake's Equation

Whilst this work does not look directly at planets and their ability to host life, it is still possible to draw some general conclusions from this work.

This work highlighted the correlation between metallicity and the number of supernovae experienced in the first 100 Myrs, with lower metallicity particles experiencing fewer supernovae in their early lives. As supernovae provide the metals required for planet formation, it is possible that lower metallicity particles would not have the material required to form planets, preventing them from hosting life. This suggests that metal poor planets are likely to reduce the value of n_e in Drake's equation.

On the other hand, metal rich particles experience a high number of supernovae in their early lives, providing them the metals needed in planet formation. However, exposure to a large number of supernovae may cause planets or their atmospheres to be damaged by photoionisation, again reducing the value of n_e. It is possible that planets that are far enough away from supernovae will not be damaged, which would allow them to potentially host life in the later stages of their evolution. As the resolution of these simulations does not allow for the distance between stars and supernovae to be calculated, it is not possible to confirm how common this is in Milky Way type galaxies.

Metal poor particles are also found to spend less time in a cluster environment. As cluster environments are densely populated with stars, which are a source of ionising radiation, they can be particularly harsh environments for planets. By limiting the amount of time spent in this environment, metal poor stars may improve the survival chances of any planets that they have been able to form, thus increasing the value of n_e. However, as previously mentioned, metal poor stars may have been unable to form planets due to a lack of the required metals.

Overall, there is likely a small time window that constitutes an 'ideal' amount of time to spend in a cluster environment. Within this time frame, protoplanetary disks will be able to utilise the metals produced by supernovae to form planets, before leaving and ensuring the planets are not subject to high levels of radiation. As the time window is likely to be relatively narrow, most stars are likely to either spend too much or too little time in a cluster environment, reducing the number of habitable planets that they host. This would suggest that the value of n_e is less than 1, but future work focusing on planet formation would be needed to further constrain this.

7.1.5 Final Conclusions

Overall, these results show repeated trends between particle metallicity and other measured variables. However, as particles with higher metallicities are often found in regions of higher stellar density (e.g Pechetti et al., 2020), it is possible that this identified trend is actually due to a link with stellar density. As these trends are seen in when several different variables are measured, it suggests that the metallicty (or stellar density) plays an important role in determining the evolutionary history of stellar particles, and, by proxy, stars.

This work also highlights the use of metallicity as a potential diagnostic tool. As the stellar density of the formation environment of a star impacts the metallicity, this property could be used as a proxy.

7.2 Future Work

As this work is one of the first pieces of large scale analysis done on the EMP simulations, there are ample opportunities for either extending this work, or utilising the EMP simulations in other ways. The sections that follow do not represent a comprehensive list, but highlight some of the possibilities more pertinant to this work.

7.2.1 Analysing the Evolutionary Histories of Individual Clusters

The EMP simulations use subgrid descriptions to model the formation and evolution of the stellar populations found in clusters (Reina-Campos et al., 2022a). The data from the EMP simulations includes information about these populations, so the most obvious first step in expanding on this work would be to analyse their evolutionary histories. As each particle can contain more than 1 cluster, this would require an increased level of computational time, bringing with it increased expense. However, the code utilised in this work can be easily adapted to carry out this analysis, and would provide a greater level of detail regarding the distribution of cluster mass, and enable more accurate calculations of expectation values.

Whilst the computational power required to spatially resolve individual clusters is prohibitive, it would also be possible to implement a prescription for cluster density profiles. This would allow the stellar density in clusters to be inferred , which in turn would allow for the distance between stars to be estimated. As the proximity to other stars can impact the formation and evolution of proto-planetary and planetary systems, determining this will aid in understanding how these systems will be impacted by their neighbours.

7.2.2 Further Evolution

In order to confirm some of the suggested conclusions presented in this work, the EMP simulations would need to be evolved into the future. This has been done previously (e.g Salcido et al., 2018; Oh et al., 2021), but, as with all simulations, it requires models and parameters to be predetermined. As the models and parameters chosen will dictate how the simulation evolves, small initial errors will compound over time, causing the evolution seen in the simulations to deviate further from the true evolution.

Unlike simulations that evolve only to the present day, simulations evolving into the future could not be compared to observational data. This removes the possibility of 'tweaking' the models used to ensure that the results match observations. This means that any results could only be confirmed by future observations that would occur 10^{6-9} years in the future.

However, if the EMP simulations were evolved in this way, it would enable the trends in disruption time found in this work to either be confirmed or proven wrong. It may also provide further insight into the potential consequences of the merger in galaxy 22.

7.2.3 Additional Comparison Stars

At the time of writing, the NASA Exoplanet Archive contains 765 stars that host 2 or more planets, and have recorded [Fe/H] and age values. Based on the cuts used in this work, 100 of these stars would be considered metal rich, 78 would be metal poor, and the rest would be of solar metallicity. Whilst this initially suggests that solar metallicty stars are more likely to host planets, it is important to acknowledge that exoplanets are much easier to detect when they are in the solar neighbourhood, which contains many more stars with metallicities similar to that of the Sun, giving rise to a bias.

Whilst many of these systems have well constrained metallicities, there can be large errors on the measured ages. This means that not all of these stars would fall in to one of the previously defined age- metallicity groups, so it may not be possible to repeat the analysis in the same way that it was repeated using TRAPPIST. However, as the current age-metallicity groups were decided upon based on the age and metallicity values for the Sun, it would be possible to use the code to redefine the age-metallicity groups based on the measurements for an alternative star.

By incorporating additional comparison stars that host multi- planet systems, it would be possible to compare the evolutionary histories of a range of stars, and therefore planet systems. This may highlight similarities in the evolutionary histories of these

stars that may indicate environments that are more likely to produce planet systems. This information could then be used to identify and prioritise regions that are more likely host exoplanets for observational purposes.

7.2.4 Increased Resolution

These results presented in this work use particles as a proxy for stars as the resolution does not allow for individual stars to be resolved. This means that the results make use of statistical measures, such as expectation values, means, and medians.

If the resolution of the simulations was increased, it would be possible to identify and analyse individual stars. This would allow metrics such as the amount of time spent in a cluster and the galactocentric radius of individual stars to be calculated, improving the accuracy of the results.

This level of resolution may also allow for the individual masses of stars to be accounted for, meaning stars that are likely to end their lives in supernovae events could be identified. It would then be possible to determine if these supernovae are near enough to other stars to destroy or disrupt any planets. This would provide a better insight into how both the number and proximity of supernovae impact planet formation and evolution.

Finally, identifying the evolutionary histories of individual stars will allow the point at which they transition from a cluster to the field to be identified. This will allow the cluster mass to be tracked over cosmic time, accurately determining when it has been fully disrupted. It may also be possible to identify periods of rapid mass loss and the events associated with them (e.g tidal stripping).

However, increasing the resolution of the EMP simulations would necessitate an increase in the number of particles, which in turn would increase the amount of memory required. In order to resolve and individual star, the particle separation of the simulation would need to be smaller than the star's radius. Taking a star like the Sun, which has a radius of 2.3x10^{-8} pc, and a box with sides of 100Mpc, the number of particles required can be calculated using Equation 7.1:

$$Particles = \left(\frac{100 \times 10^{6} pc}{2.3 \times 10^{-8} pc}\right)^{3} = 8.2 \times 10^{46} \tag{7.1}$$

The EAGLE simulations utilised 1.2×10^{9} particles for a box of this size (Schaye et al., 2015), meaning that it is 37 orders of magnitude too small to resolve an individual star. This makes increasing the resolution to this level an unfeasible option.

7.2.5 Computational and Simulation Adaptations

7.2.5.1 Improved Sub-Grid Routines

Due to limitations on the resolution, simulations utilise sub-grid routines to model processes that occur on scales that cannot be modelled (see Section 2.7.2 for details). Improving these routines can help to improve the accuracy of the simulations, whilst not sufficiently increasing the cost of running the simulations.

For example, the sub-grid routine used to model supernovae feedback in the EMP simulations only include thermal heating, but those included in the FIRE-3 simulations include both kinetic and thermal energy (Hopkins et al., 2023). The inclusion of kinetic energy would help to accurately model galactic winds that can expel gas and limit the material available for star formation. These gas outflows can also carry heavy elements into the intergalactic medium, which contributes to the chemical evolution of the universe.

A major advantage of sub-grid routines is that they are rarely computationally intensive, meaning they would not add significantly to the computational time or cost of running the simulation. However, developing a sub-grid routine can be time consuming, and as advances are made in our understanding of the universe, they will require improving or replacing.

7.2.5.2 Graphics Processing Units

Graphics Processing Units (GPUs) were designed to render the graphics commonly seen in video games (Hirling et al., 2023). In order to do this, they were designed with thousands of small cores, which enable them to simultaneously run multiple tasks. This is in stark contrast to central processing units (CPUs) which consist of fewer, more powerful cores, that are designed to process tasks sequentially. The core design in GPUs make them ideal for carrying out many calculations simultaneously, meaning that if the complex physics found in cosmological simulations can be broken down into smaller calculations, GPUs can carry out the process much faster than CPUs.

This is of particular use in hydrodynamical simulations, such as the EMP simulations, as these often involve complex physics that can be broken down into smaller, independent calculations, making it easy to parallelise (Hirling et al., 2023). These equations are therefore solved much faster on a GPU, which is designed to carry out calculations in this way (Sokolowski et al., 2024). This also makes GPUs cheaper than CPUs, as the

reduced run time is associated with a reduced power consumption, and therefore lower costs (Sokolowski et al., 2024).

However, as GPUs are a relatively new introduction to cosmological simulations, much of the code available today has been written to run on CPUs. Adapting this code to enable it to run effectively on GPUs is time consuming and complex, meaning it is unlikely that old code will be adapted to work in this way.

7.2.5.3 Emulators

Another potential adaptation would be the use of emulators, which are capable of approximating the outputs of traditional simulations. Current emulators are typically neural networks that have found the most efficient way to model the required physics (Kasim et al., 2022). Once this has been identified, the emulator is trained via supervised learning. The training data consists of input-output pairs from the full simulation the emulator is being used to model. Based on input data (e.g cosmological parameters or initial conditions) the neural network learns to predict outputs such as the matter distribution in the universe (Kasim et al., 2022).

Once the model has been trained, it is presented with a new set of unseen input data (Kasim et al., 2022). The model is then evaluated based on how accurately it reproduces the output associated with each input, ensuring the model is able to adapt to unseen data.

The final neural network is able to predict the outcome of the simulation when presented with new input parameters. If a simulation would take several days to produce an output, a neural network emulator is capable of producing it in a few seconds if run on a CPU, or even quicker if a GPU is utilised (Kasim et al., 2022).

Emulators designed in this way can be applied to simulations on various scales, from modelling sub-grid physics, to predicting the outcomes of full simulation runs, making them incredibly diverse. They also drastically speed up the process of obtaining results when individual inputs are changed, allowing the most accurate set of input parameters to be quickly determined. However, the initial determination of the most efficient modelling method can be computationally expensive, and a range of high quality data is required for both training and testing purposes. This often means that the initial cost of creating these emulators is high, but this is then offset by the reduction in computational cost when it is deployed.

Bibliography

Agol E., et al., 2021, The Planetary Science Journal, 2, 1

Ahrer T. J. T. E. C. E. R., et al., 2022, Nature, 614, 649

Akeson R. L., et al., 2013, , 125, 989

Alexander R., 2014, in Booth M., Matthews B. C., Graham J. R., eds, IAU Symposium Vol. 299, Exploring the Formation and Evolution of Planetary Systems. pp 179–189 (arXiv:1308.1791), doi:10.1017/S1743921313008296

Alexander P. E. R., Gieles M., 2012, Monthly Notices of the Royal Astronomical Society, 422, 3415

Amard L., Matt S. P., 2020, The Astrophysical Journal, 889, 108

Anders, F. et al., 2017, A&A, 600, A70

Babcock H. W., 1939, Lick Observatory Bulletin, 498, 41

Baldry I. K., et al., 2012, Monthly Notices of the Royal Astronomical Society, 421, 621

Bastian N., Pfeffer J., Kruijssen J. M. D., Crain R. A., Trujillo-Gomez S., Reina-Campos M., 2020, Monthly Notices of the Royal Astronomical Society, 498, 1050

Bell E. F., de Jong R. S., 2000, Monthly Notices of the Royal Astronomical Society, 312, 497

Belokurov V., Kravtsov A., 2022, Monthly Notices of the Royal Astronomical Society, 514, 689

Bonanno A., Schlattl H., Paternò L., 2002, Astronomy & Astrophysics, 390, 1115–1118

Bournaud F., Jog C. J., Combes F., 2007, Astronomy & Astrophysics, 476, 1179

Burgasser A. J., Mamajek E. E., 2017, The Astrophysical Journal, 845, 110

Burke C. J., et al., 2014, , 210, 19

Burkhart B., 2018, The Astrophysical Journal, 863, 118

Cabrera J., et al., 2014, The Astrophysical Journal, 781, 18

Cassan A., et al., 2012, , 481, 167

Chabrier G., 2003, , 115, 763

Chabrier G., 2005, in Corbelli E., Palla F., Zinnecker H., eds, Astrophysics and Space Science Library Vol. 327, The Initial Mass Function 50 Years Later. p. 41 (arXiv:astro-ph/0409465), doi:10.1007/978-1-4020-3407-7˙5

Chevance M., et al., 2020, , 216, 50

Choksi N., Kruijssen J. M. D., 2021, Monthly Notices of the Royal Astronomical Society, 507, 5492

Ćirković M. M., 2004, Astrobiology, 4, 225

Clement M. S., Izidoro A., Raymond S. N., Deienno R., 2024, arXiv e-prints, p. arXiv:2411.03453

Close J. L., Pittard J. M., 2017, Monthly Notices of the Royal Astronomical Society, 469, 1117

Crain R. A., van de Voort F., 2023, Annual Review of Astronomy and Astrophysics, 61, 473–515

Crain R. A., et al., 2015, Monthly Notices of the Royal Astronomical Society, 450, 1937

Crain R. A., et al., 2017, Monthly Notices of the Royal Astronomical Society, 464, 4204

D'Angelo G., Durisen R. H., Lissauer J. J., 2010, Giant Planet Formation (arXiv:1006.5486)

Dalla Vecchia C., Schaye J., 2012, Monthly Notices of the Royal Astronomical Society, 426, 140

Davies P., 2007, Scientific American, 297, 62

Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, The Astrophysical Journal, 292, 371

Davis A. M., Alexander C. M. O. D., Ciesla F. J., Gounelle M., Krot A. N., Petaev M. I., Stephan T., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, Protostars and Planets VI. pp 809–831, doi:10.2458/azu˙uapress˙9780816531240-ch035

De Lucia G., Springel V., White S. D. M., Croton D., Kauffmann G., 2006, Monthly Notices of the Royal Astronomical Society, 366, 499

De Rossi M. E., Bower R. G., Font A. S., Schaye J., Theuns T., 2017, Monthly Notices of the Royal Astronomical Society, 472, 3354

Deason A. J., Belokurov V., 2024, Galactic Archaeology with Gaia (`arXiv:2402.12443`)

Doner S., Ak S., Tas O. O., Plevne O., 2023, The Age-Metallicity Relation in the Solar Neighbourhood (`arXiv:2304.14747`)

Donlon T., Newberg H. J., Sanderson R., Bregou E., Horta D., Arora A., Panithanpaisal N., 2024, Monthly Notices of the Royal Astronomical Society, 531, 1422

Drake F. D., 1961a, Physics Today, 14, 40

Drake F. D., 1961b, Physics Today, 14, 40

Drazkowska J., et al., 2023, Planet Formation Theory in the Era of ALMA and Kepler: from Pebbles to Exoplanets (`arXiv:2203.09759`)

Dye S. T., 2012, Reviews of Geophysics, 50, RG3007

Eggen O. J., Lynden-Bell D., Sandage A. R., 1962, The Astrophysical Journal, 136, 748

Ellis J., Schramm D. N., 1995, Proceedings of the National Academy of Science, 92, 235

Elmegreen B. G., 2008, The Astrophysical Journal, 672, 1006

Enya K., Yamagishi A., Kobayashi K., Yoshimura Y., Tasker E. J., 2023, Astrobiology, 23, 1099

Fischer D. A., Valenti J., 2005, The Astrophysical Journal, 622, 1102

Font A. S., McCarthy I. G., Johnstone D., Ballantyne D. R., 2004, The Astrophysical Journal, 607, 890

Freeman K. C., 1970, The Astrophysical Journal, 160, 811

Furlong M., et al., 2015, Monthly Notices of the Royal Astronomical Society, 450, 4486

Furlong M., et al., 2016, Monthly Notices of the Royal Astronomical Society, 465, 722

Gillon M., et al., 2017, , 542, 456

Goddard Q. E., Bastian N., Kennicutt R. C., 2010, Monthly Notices of the Royal Astronomical Society, 405, 857

Goderis S., Chakrabarti R., Debaille V., Kodolányi J., 2016, J. Anal. At. Spectrom., 31, 841

Gonzalez G., 2005, Origins of Life and Evolution of the Biosphere, 35, 555

Gounelle M., Chaussidon M., Rollion-Bard C., 2013, , 763, L33

Hartmann L., Herczeg G., Calvet N., 2016, Annual Review of Astronomy and Astrophysics, 54, 135

Hennebelle P., Chabrier G., 2011, , 743, L29

Heyer M., Krawczyk C., Duval J., Jackson J. M., 2009, The Astrophysical Journal, 699, 1092

Hirling P., Bianco M., Giri S. K., Iliev I. T., Mellema G., Kneib J.-P., 2023, arXiv e-prints, p. arXiv:2311.01492

Holtzman J. A., et al., 1992, , 103, 691

Hopkins P. F., et al., 2023, Monthly Notices of the Royal Astronomical Society, 519, 3154

Hughes M. E., Pfeffer J. L., Bastian N., Martig M., Kruijssen J. M. D., Crain R. A., Reina-Campos M., Trujillo-Gomez S., 2022, Monthly Notices of the Royal Astronomical Society, 510, 6190

Hunter D., 1997, , 109, 937

Inaba S., Wetherill G. W., Ikoma M., 2003, , 166, 46

Jabbari B., 1997, Proceedings of the IEEE, 85, 1523

Jeans J. H., 1902, Philosophical Transactions of the Royal Society of London Series A, 199, 1

Jenkins A., 2010, Monthly Notices of the Royal Astronomical Society, 403, 1859

Jenkins A., 2013, Monthly Notices of the Royal Astronomical Society, 434, 2094

Kaib N. A., 2018, in Deeg H. J., Belmonte J. A., eds, , Handbook of Exoplanets. p. 59, doi:10.1007/978-3-319-55333-7˙59

Kant I., 1755, Allgemeine Naturgeschichte und Theorie des Himmels

Kapteyn J. C., 1922, The Astrophysical Journal, 55, 302

Kasim M. F., et al., 2022, Machine Learning: Science and Technology, 3, 015013

Keller M. D., 1989, Biological Oceanography, 6, 375

Kellermann K. I., 2023, arXiv e-prints, p. arXiv:2302.06446

Kirby E. N., Boylan-Kolchin M., Cohen J. G., Geha M., Bullock J. S., Kaplinghat M., 2013, The Astrophysical Journal, 770, 16

Kley W., 2019, Saas-Fee Advanced Course, 45, 151

Kokubo E., Ida S., 2002, The Astrophysical Journal, 581, 666

Kopparapu Wolf M., 2005. University of Arizona Press, doi:10.2458/azu˙uapress˙9780816540068, http://dx.doi.org/10.2458/azu_uapress_9780816540068

Kratter K. M., Matzner C. D., 2006, Monthly Notices of the Royal Astronomical Society, 373, 1563

Krause, Martin G. H. Charbonnel, Corinne Bastian, Nate Diehl, Roland 2016, A&A, 587, A53

Kruijssen J. M. D., 2011, in Stellar Clusters & Associations: A RIA Workshop on Gaia. pp 137–141 (arXiv:1107.2114), doi:10.48550/arXiv.1107.2114

Kruijssen J. M. D., 2012a, Monthly Notices of the Royal Astronomical Society, 426, 3008

Kruijssen J. M. D., 2012b, Monthly Notices of the Royal Astronomical Society, 426, 3008

Kruijssen J. M. D., Pelupessy F. I., Lamers H. J. G. L. M., Portegies Zwart S. F., Icke V., 2011, Monthly Notices of the Royal Astronomical Society, 414, 1339

Kruijssen J. M. D., Maschberger T., Moeckel N., Clarke C. J., Bastian N., Bonnell I. A., 2012a, Monthly Notices of the Royal Astronomical Society, 419, 841

Kruijssen J. M. D., Pelupessy F. I., Lamers H. J. G. L. M., Portegies Zwart S. F., Bastian N., Icke V., 2012b, Monthly Notices of the Royal Astronomical Society, 421, 1927

Kruijssen J. M. D., Pfeffer J. L., Crain R. A., Bastian N., 2019a, Monthly Notices of the Royal Astronomical Society, 486, 3134

Kruijssen J. M. D., Pfeffer J. L., Reina-Campos M., Crain R. A., Bastian N., 2019b, Monthly Notices of the Royal Astronomical Society, 486, 3180

Kruijssen J. M. D., et al., 2020, Monthly Notices of the Royal Astronomical Society, 498, 2472

Krumholz M. R., 2015, in Vink J. S., ed., Astrophysics and Space Science Library Vol. 412, Very Massive Stars in the Local Universe. p. 43 (arXiv:1403.3417), doi:10.1007/978-3-319-09596-7˙3

Lacey C., Cole S., 1993, Monthly Notices of the Royal Astronomical Society, 262, 627

Lada C. J., Lada E. A., 2003, , 41, 57

Lambrechts M., Johansen A., 2012, Astronomy & Astrophysics, 544, A32

Lamers, H. J. G. L. M. Kruijssen, J. M. D. Bastian, N. Rejkuba, M. Hilker, M. Kissler-Patig, M. 2017, A&A, 606, A85

Langer N., 2012, , 50, 107

Larson R. B., 1969, Monthly Notices of the Royal Astronomical Society, 145, 271

Laughlin G., Adams F. C., 1998, , 508, L171

Lee T., Papanastassiou D. A., Wasserburg G. J., 1976, , 3, 41

Leitherer C., et al., 1999, , 123, 3

Leroy A. K., Bigiel F., Hughes A., Schinnerer E., Usero A., Usero 2016, in Jablonka P., André P., van der Tak F., eds, Vol. 315, From Interstellar Clouds to Star-Forming Galaxies: Universal Processes?. pp 175–182, doi:10.1017/S174392131600747X

Lineweaver C. H., Chopra A., 2012, Annual Review of Earth and Planetary Sciences, 40, 597

Lineweaver C. H., Fenner Y., Gibson B. K., 2004, Science, 303, 59–62

Madhusudhan N., Sarkar S., Constantinou S., Holmberg M., Piette A. A. A., Moses J. I., 2023, Carbon-bearing Molecules in a Possible Hycean Atmosphere (arXiv:2309.05566)

Maraboli E., Mantegazza F., Lodato G., 2023, European Physical Journal Plus, 138, 152

Mayr E., 1995, Bioastronomy News, 7, 2

McKee C. F., Ostriker E. C., 2007, , 45, 565

McKee C. F., Tan J. C., 2003, The Astrophysical Journal, 585, 850

McKinney M. L., 1989, Science, 243, 103

Melott A. L., Thomas B. C., Kachelrieß M., Semikoz D. V., Overholt A. C., 2017, The Astrophysical Journal, 840, 105

Meynet G., Mowlavi N., Maeder A., 2006, Massive star evolution at high metallicity (arXiv:astro-ph/0611261)

Montmerle T., Augereau J.-C., Chaussidon M., Gounelle M., Marty B., Morbidelli A., 2006, Earth Moon and Planets, 98, 39

Moster B. P., Naab T., White S. D. M., 2013, Monthly Notices of the Royal Astronomical Society, 428, 3121

2009, Nature, 461, 316–316

Oh B. K., Peacock J. A., Khochfar S., Smith B. D., 2021, Monthly Notices of the Royal Astronomical Society, 507, 5432

Ott U., 2016, Isotope Variations in the Solar System: Supernova Fingerprints. Springer International Publishing, Cham, pp 1–27, doi:10.1007/978-3-319-20794-0˙17-1, https://doi.org/10.1007/978-3-319-20794-0_17-1

Papaloizou J. C. B., Nelson R. P., Kley W., Masset F. S., Artymowicz P., 2006, Disk-Planet Interactions During Planet Formation (arXiv:astro-ph/0603196)

Pechetti R., Seth A., Neumayer N., Georgiev I., Kacharov N., den Brok M., 2020, The Astrophysical Journal, 900, 32

Petigura E. A., Howard A. W., Marcy G. W., 2013, Proceedings of the National Academy of Sciences, 110, 19273–19278

Pfalzner, S. 2013, A&A, 549, A82

Pfeffer J., Kruijssen J. M. D., Crain R. A., Bastian N., 2017, Monthly Notices of the Royal Astronomical Society, 475, 4309–4346

Pfeffer J., Kruijssen J. M. D., Crain R. A., Bastian N., 2018, Monthly Notices of the Royal Astronomical Society, 475, 4309

Planck Collaboration et al., 2014a, Astronomy & Astrophysics, 571, A1

Planck Collaboration et al., 2014b, Astronomy & Astrophysics, 571, A16

Planck Collaboration et al., 2016, Astronomy & Astrophysics, 594, A13

Pollack J. B., Hubickyj O., Bodenheimer P., Lissauer J. J., Podolak M., Greenzweig Y., 1996, , 124, 62

Pontzen A., Roškar R., Stinson G., Woods R., 2013, pynbody: N-Body/SPH analysis for python, Astrophysics Source Code Library, record ascl:1305.002

Portegies Zwart, S. Pelupessy, I. van Elteren, A. Wijnen, T. P. G. Lugaro, M. 2018, A&A, 616, A85

Portegies Zwart S. F., McMillan S. L. W., Gieles M., 2010, , 48, 431

Rahmati A., Schaye J., Crain R. A., Oppenheimer B. D., Schaller M., Theuns T., 2016, Monthly Notices of the Royal Astronomical Society, 459, 310

Raymond S. N., Morbidelli A., 2022, Planet Formation: Key Mechanisms and Global Models. Springer International Publishing, p. 3–82, doi:10.1007/978-3-030-88124-5˙1, http://dx.doi.org/10.1007/978-3-030-88124-5_1

Raymond S. N., Quinn T., Lunine J. I., 2006, Icarus, 183, 265–282

Raymond S. N., Quinn T., Lunine J. I., 2007, Astrobiology, 7, 66–84

Reina-Campos M., Kruijssen J. M. D., 2017, Monthly Notices of the Royal Astronomical Society, 469, 1282

Reina-Campos M., Keller B. W., Kruijssen J. M. D., Gensior J., Trujillo-Gomez S., Jeffreson S. M. R., Pfeffer J. L., Sills A., 2022a, arXiv e-prints, p. arXiv:2202.06961

Reina-Campos M., Trujillo-Gomez S., Deason A. J., Kruijssen J. M. D., Pfeffer J. L., Crain R. A., Bastian N., Hughes M. E., 2022b, Monthly Notices of the Royal Astronomical Society, 513, 3925

Robitaille T. P., Whitney B. A., 2010, , 710, L11

Rubin V. C., Ford W. Kent J., 1970, The Astrophysical Journal, 159, 379

Rubin V. C., Ford W. K. J., Thonnard N., 1978, , 225, L107

Salcido J., et al., 2018, Monthly Notices of the Royal Astronomical Society, 477, 3744

Schaye J., 2004, The Astrophysical Journal, 609, 667

Schaye J., Dalla Vecchia C., 2008, Monthly Notices of the Royal Astronomical Society, 383, 1210

Schaye J., et al., 2015, Monthly Notices of the Royal Astronomical Society, 446, 521

Schechter P., 1976, The Astrophysical Journal, 203, 297

Schinnerer E., Leroy A. K., 2024, arXiv e-prints, p. arXiv:2403.19843

Searle L., Zinn R., 1978, The Astrophysical Journal, 225, 357

Sellwood J. A., Masters K. L., 2022, , 60

Shallue C. J., Vanderburg A., 2018, , 155, 94

Shields G. A., 1974, The Astrophysical Journal, 193, 335

Sohn S. T., Anderson J., van der Marel R. P., 2012, The Astrophysical Journal, 753, 7

Sokolowski M., Aniruddha G., Di Pietrantonio C., Harris C., Price D. C., McSweeney S., Wayth R. B., Bhat N. D. R., 2024, arXiv e-prints, p. arXiv:2405.13478

Springel V., 2005, Monthly Notices of the Royal Astronomical Society, 364, 1105

Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, Monthly Notices of the Royal Astronomical Society, 328, 726

Spurzem R., Giersz M., Heggie D. C., Lin D. N. C., 2009, The Astrophysical Journal, 697, 458

Süli Á., 2010, in Journal of Physics Conference Series. p. 012004, doi:10.1088/1742-6596/218/1/012004

Swain M. R., Hasegawa Y., Thorngren D. P., Roudier G. M., 2024, Space Science Reviews, 220

Tan J. C., Beltrán M. T., Caselli P., Fontani F., Fuente A., Krumholz M. R., McKee C. F., Stolte A., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, Protostars and Planets VI. pp 149–172 (arXiv:1402.0919), doi:10.2458/azu˙uapress˙9780816531240-ch007

Thommes E. W., Duncan M. J., Levison H. F., 1999, , 402, 635

Thommes E., Duncan M., Levison H., 2003, Icarus, 161, 431–455

Trayford J. W., et al., 2015, Monthly Notices of the Royal Astronomical Society, 452, 2879

Trujillo-Gomez S., Reina-Campos M., Kruijssen J. M. D., 2019, Monthly Notices of the Royal Astronomical Society, 488, 3972

Trujillo-Gomez S., DiederikÂ Kruijssen J. M., Reina-Campos M., Pfeffer J. L., Keller B. W., Crain R. A., Bastian N., Hughes M. E., 2021, Monthly Notices of the Royal Astronomical Society, 503, 31

Tsumura K., 2020, Scientific Reports, 10, 12795

Usher C., Pfeffer J., Bastian N., Kruijssen J. M. D., Crain R. A., Reina-Campos M., 2018, Monthly Notices of the Royal Astronomical Society, 480, 3279

Ventura P., Dell'Agli F., Lugaro M., Romano D., Tailo M., Yague A., 2020, Gas and dust from metal-rich AGB stars (`arXiv:2007.02120`)

Vinn O., 2024, Science Progress, 107, 00368504241272491

Webb J. J., Reina-Campos M., Kruijssen J. M. D., 2019, Monthly Notices of the Royal Astronomical Society, 486, 5879

Westby T., Conselice C. J., 2020, The Astrophysical Journal, 896, 58

White S. D. M., Rees M. J., 1978, Monthly Notices of the Royal Astronomical Society, 183, 341

Wiersma R. P. C., Schaye J., Smith B. D., 2009a, Monthly Notices of the Royal Astronomical Society, 393, 99

Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009b, Monthly Notices of the Royal Astronomical Society, 399, 574

Woosley S. E., Weaver T. A., 1995, , 101, 181

Zinnecker H., Yorke H. W., 2007, , 45, 481

Zwicky F., 1933, Helvetica Physica Acta, 6, 110

de Wit J., et al., 2018, Nature Astronomy, 2, 214–219

van der Marel R. P., Besla G., Cox T. J., Sohn S. T., Anderson J., 2012, The Astrophysical Journal, 753, 9